Discovery of novel pathways of microbial arginine biosynthesis

Thesis by

Juan Manuel Cabrera Luque

CHILDREN'S NATIONAL MEDICAL CENTER
Research Center for Genetic Medicine
Washington DC, USA
2010

Principal Investigator Hiroki Morizono, Ph.D. by the University of Minnesota, Department of Integrative Systems Biology, Children's National Medical Center

CERTIFIES, that the present work "DISCOVERY OF NOVEL PATHWAYS OF MICROBIAL ARGININE BIOSYNTHESIS" has been developed by Juan Manuel Cabrera Luque under his supervision in the Department of Integrative Systems Biology, Children's National Medical Center in Washington DC as a thesis project in order to obtain a Ph.D. in Biochemistry and Molecular Biology at University of Valencia, Biochemistry and Molecular Biology Department.

Hiroki Morizono, Ph.D.

A mis hijas María y Beatriz.
A mi Gema, a mi hermano Raúl
y a mis queridos padres Juan y Conchi

"In science, an experiment is as good as its controls."

Mendel Tuchman

Acknowledgements:

There are a great number of people that I need to thank for providing me with the professional, technical and emotional support that I so genuinely appreciated and without which the completion of this work would not have been possible.

To my mentor, Hiroki Morizono, Ph.D. I would like to express my sincere gratitude for your support, patience and encouragement over the years. I have learned a great deal from you and I thank you for the guidance and for the enthusiasm for science that you have instilled in me.

To Ljubica Caldovic, PhD., I want to thank you for being my "second thesis mentor" and for the generosity of your time and help, which have resulted in a valuable contribution to this work.

To Mendel Tuchman, M.D., I want to thank you for your support and for giving me the opportunity to work in your lab in the urea cycle disorders field. I have learned from you more things that I can express in a few pages, for that and for many more reasons, thank you.

To Michael Malamy, PhD., who first identified the essential genes for the arginine biosynthetic pathway in *B. fragilis* and who made possible the initial stages of this project, thank you.

To my friends and colleagues, Net Haskins and Himani Manjumdar. I have learned so much from you and I have appreciated all your help and support. For all the discussions, "lab lunches" and all the hours we have spend together trying to solve the mysteries of the urea cycle, protein interactions and Western blot results… It has been a pleasure working with you and I am lucky that our paths have crossed.

To Lisa Sheehy and Marie Pichaske for all your help throughout these years and the much paperwork you helped me fill out, I owe you a great thank you,

To all the GenMed people who have been there throughout all these years making them some of the most memorable ones for me. I owe a debt of gratitude to Kristy Brown, Javad Nazarian, Gustavo Nader, Yetrib Hathout, Joe Devaney, Mary Gallo and many more for those useful conversations that allowed me to disconnect from time to time.

To Vicente Rubio, PhD for your useful comments and review of the initial version of this thesis, thank you.

To Gary Cunningham I want to thank you for your time and your valuable reviews of my English writing.

To Marshall Summar, M.D., for your support in the last steps of this thesis, whithout your generosity, understanding and "time" it would have been really hard to finalyze. THANK YOU!

Table of Contents

Abbreviations

AAK: amino acid kinase
ADI: arginine deiminase pathway
ANorv: N-acetyl-L-norvaline
AO: acetylornithinase
AOD: N-acetyl-L-ornithine deacetylase
AOrn: αN-acetylornithine
AOTC: N-acetyl-L-ornithine transcarbamylase
Arg: L-arginine
AS: argininosuccinate
ASS: argininosuccinate synthase
ASer: acetyl-L-serine
ASL: argininosuccinate lyase
ATC: aspartate transcarbamylase
BSA: bovine serum albumin
cOTC: catabolic OTC
Cit: L-citrulline
CK: carbamate kinase
CP: carbamylphosphate
EDTA: ethylenediaminetetraacetic acid
Gln: glutamine
Glu: L-glutamate
LC-MS: liquid chromatography coupled with mass spectrometry
LUCA: Last Universal Common Ancestor
NAG: N-acetyl-L-glutamate
NAGK: N-acetyl-L-glutamate kinase
NAGS: N-acetyl-L-glutamate synthase
NCBI: National Center for Biotechnology Information
NO: nitric oxide
Norv: L-norvaline
NSG: N-succinyl glutamate
NSGK: N-succinyl glutamate kinase
n/a: non applicable
OAT: ornithine acetyltransferase
Orn: L-ornithine
OTC: ornithine transcarbamylase
PALA: N^{δ}-phosphonoacetyl-L-aspartate
PALAO: αN-acetyl-δN-phosphonoacetyl-L-ornithine
PALO: N^{δ}-phosphonoacetyl-L-ornithine
PALSO: αN-succinyl-δN-phosphonoacetyl-L-ornithine
PP: pyrophosphate
PTC: putrescine transcarbamylase
SCit: N-succinyl-L-citrulline
Ser: L-serine
SOTC: N-succinyl-L-ornithine transcarbamylase
SOrn: N-succinyl-L-ornithine
S-NAGS: short NAGS
TCA: Tri-chloroacetic acid
UTC: unknown transcarbamylase

Chapter 1: General introduction and purpose of this work

1.1 The importance of arginine and of its *de novo* synthesis

The amino acid L-arginine is an essential component of all living organisms. Its importance resides in the variety of functions that arginine itself, along with some intermediary metabolites involved in its *de novo* synthesis in the cell. Figure 1 shows the main metabolic interconnections of the *de novo* biosynthesis of arginine and its intermediates.

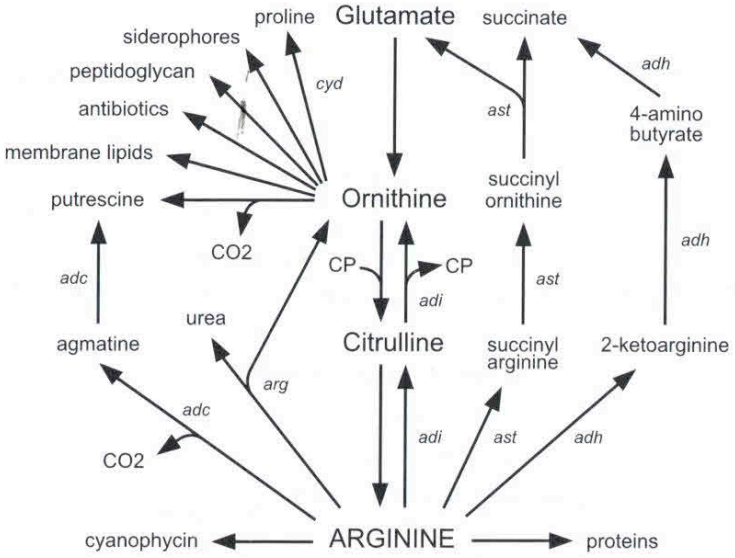


Figure 1: Schematic representation of the main metabolic interconnections of Arg biosynthesis. Adapted from[1]. Only key intermediates are shown. adi: Arg deiminase pathway; ast: Arg succinylase pathway; adh: Arg oxidase/dehydrogenase/transaminase pathway[2]; arg: arginase pathway; adc: Arg decarboxylase pathway; cyd: Orn cyclodeaminase; CP: carbamylphosphate.

In addition to the above-mentioned important metabolic interconnections, Arg has the following essential roles:

- used as building block of many proteins
- precursor of the energy storage compounds:
    - arginine phosphate, involved in muscle contraction in invertebrates, and
    - creatine phosphate, which is stored in muscle and brain in vertebrates.
- intermediate metabolite of the urea cycle.
- precursor of the signaling molecules nitric oxide (NO), agmatine and glutamate in mammals.
- nitrogen storage compound in plants and cyanobacteria

In many prokaryotes, fungi, and plants the *de novo* biosynthesis of Arg proceeds from glutamate in eight enzymatic steps (Figure 2). The first committed step of this pathway is the N-acetylation of glutamate. Acetylation of the early

precursors of Arg distinguishes them from the analogous intermediates in the biosynthesis of proline[3, 4].

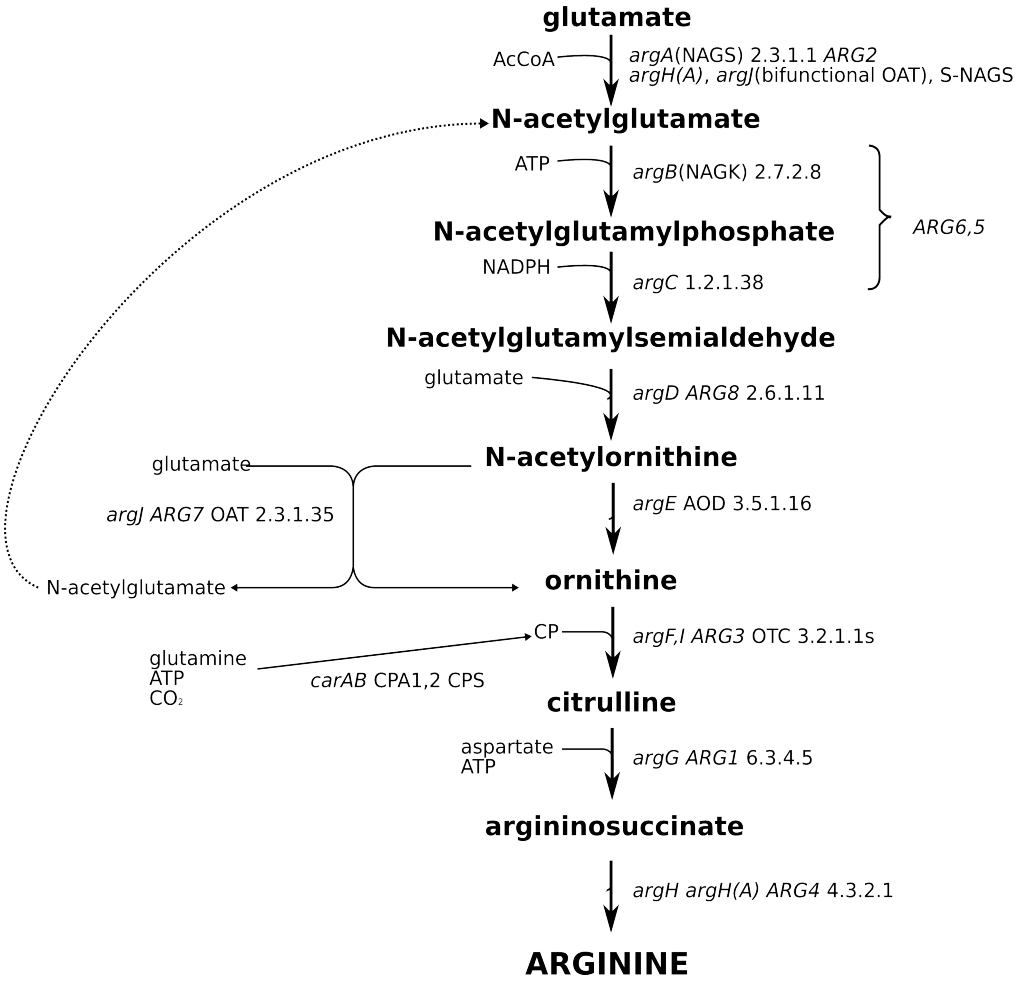


Figure 2: Outline of pathways for the *de novo* arginine biosynthesis from L-glutamate. Prokaryotic genetic symbols are in small italics, yeast genetic symbols in capital italics. EC numbers are indicated.

The Arg biosynthesis pathway in prokaryotes and lower eukaryotes and the mammalian urea cycle share several enzymes and intermediates, leading to the view that the urea cycle has evolved from the Arg biosynthesis pathway. In addition, enzymes of the urea cycle and their counterparts in the Arg biosynthetic pathway share sequence and structural similarities.

One of the key steps, the synthesis of the intermediate metabolite ornithine can be catalyzed by either of two enzymes (Figure 2) from αN-acetylornithine (AOrn):

(i) AOrn deacetylase (AOD, EC 3.5.1.16), also called acetylornithinase (AO), or,
(ii) ornithine acetyltransferase (OAT, EC 2.3.1.35).

Epitomizing the linear and "cyclic" pathways respectively.

1.2 The linear pathway of arginine biosynthesis

In the linear arginine biosynthetic pathway, the flow of acetylated precursors, which, in most cases, starts with the acetylation of glutamate by the enzyme N-acetyl-L-glutamate synthase (NAGS), leads to the formation of AOrn (Figure 2). AOrn is then deacetylated to form ornithine by the enzyme acetylornithine deacetylase (AOD) the product of the gene ArgE (in bacteria).

In the last part of the pathway, Arg is synthesized in three additional steps from ornithine. The first of them involves the carbamylation of Orn to form L-citrulline (Cit). In most organisms studied until now, Orn is carbamylated to Cit by the enzyme ornithine transcarbamylase (OTC, EC 2.1.3.3). *Escherichia coli* was found to contain two paralogous genes, *argF* and *argI* whose products can form hybrid trimeric forms of OTC[5].

The second substrate of the reaction catalyzed by OTC is carbamylphosphate (CP). In most microorganisms CP is synthesized from HCO_3^-, ATP, and glutamine by the enzyme carbamylphosphate synthetase (CPS, EC 6.3.5.5). CP can be synthesized by a single CPS or by two distinct enzymes. The single-CPS situation is typical of Gram-negative bacteria such as *E. coli* and *Pseudomonas aeruginosa*[6, 7]. The double-CPS situation has been found in some Gram-positive bacteria, represented by *Bacillus subtilis*, and in fungi[8]. In those organisms devoid of CPS, CP may be generated from Arg by the catabolic arginine deiminase (ADI) pathway[9], or, perhaps in the hyperthermophilic archaeon *Pyrococcus furiosus*[10, 11] by a carbamate kinase using ATP and carbamate as substrates.

The intermediate citrulline is converted by the enzyme argininosuccinate synthase (ASS, EC 6.3.4.5) into argininosuccinate (AS), which is then split into Arg and fumarate by the enzyme argininosuccinate lyase (ASL, EC 4.3.2.1) in the last step of the Arg biosynthetic pathway (Figure 2).

1.3 Variant of the classical pathway of arginine biosynthesis and its occurrence

Glutamate can be acetylated in the presence of OAT from AOrn in the "cyclic" pathway (Figure 2), to form Orn and N-acetyl-L-glutamate (NAG). This enzyme was discovered in *Micrococcus glutamicus* as a substitute for AOD[12] while studying the fermentation of Orn in this microorganism.

In some cases, OAT is able to synthesize NAG *de novo* from AcCoA and glutamate, however, there are no detectable similarities between these bifunctional OATs and the classical NAGS[13]. Since its discovery in the late 50's, OAT has been found encoded in the genome of 3531 organisms (Table 1). In most cases the corresponding gene product have been annotated as bifunctional (that is, transacetylase as well as NAGS) on the sole basis of sequence similarity, but only a small subset of these gene products have been experimentally characterized to confirm whether they are indeed mono- or bifunctional OATs[13].

The presence of a mono- or bifunctional OAT does not necessarily exclude the presence of NAGS, since in some instances both genes are present in the same microorganism[13]. In those organisms with a monofunctional OAT, NAGS is supposed to play an anaplerotic role by priming the acetyl cycle with the first molecule of NAG. Most intriguing is the functional redundancy in those organisms with bifunctional OAT, which also contain NAGS; In these organisms the AOrn-dependent transacetylation is believed to be quantitatively the most important reaction for NAG production[13], although no data has been reported that confirm this hypothesis.

Table 1: Occurrence of OAT in sequenced genomes. The sequences in the database at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/protein) were searched using "ornithine acetyltransferase" as query. The first figure is the number of hits, whereas the second number in each pair is the number of sequenced genomes at the time of the search.

Bacteria (3148/173101)		
	Proteobacteria (1378/75698)	
		α-proteobacteria (458)
		β-proteobacteria (433)
		γ-proteobacteria (340)
		Others (147)
	Firmicutes (848/26944)	
		Bacillales (477)
		Others (371)
	Actinobacteria (524/27861)	
		Actinomycetales (460)
		Others (64)
	Cyanobacteria (136/6916)	
	GNS bacteria (42)	
	Spirochaetes (27/1446)	
	Green sulfur bacteria (26)	
	Others (167)	
Eukaryotes (291/291145)		
	Fungi (205/62076)	
	Green plants (65/113035)	
	Others (21)	
Archaea (88/5220)		
	Euryarchaeotes (87/3201)	
	Others (1)	
Unclassified (2/55)		
Viruses (2/63967)		

This "cyclic" pathway has been referred frequently to as "more evolved" than the linear pathway, but biochemical and phylogenetic analyses suggest that the loss of OAT could have been one of the reasons for the recruitment of AOD in the pathway of arginine biosynthesis[13].

The fact that there are a few organisms in which the NAGS gene is absent while an apparently bifunctional OAT is present leads to the conclusion that these organisms rely on a bifunctional OAT for NAG synthesis and therefore for Arg biosynthesis.

Recent studies have revealed that there are other alternatives to OAT for the catalysis of the first committed step of the arginine biosynthesis pathway.

1.4 The catalyst of the first committed step of the pathway

As stated before, the first committed step of the *de novo* Arg biosynthetic pathway is the αN-acetylation of L-glutamate. In prokaryotes, several enzymes can catalyze this reaction:

(i) Classical NAGS, first characterized in *E. coli*, and present also in those animals that synthesize arginine.

(ii) Bifunctional OAT, able to use AcCoA to acetylate glutamate and also utilizing AOrn for transacetylation of glutamate. Present in Bacteria, Archaea and many Eukaryotes (Table 1).

(iii) Short NAGS of the GCN5-related N-acetyltransferase (GNAT) family, either independent or fused with the ArgH protein (argininosuccinase)

The first two enzymes on this list have already been discussed in the preceding sections.

Genetic studies of *Campylobacter jejuni* and *Moritella abyssi* and *Moritella profunda* provided the first evidences of the existence of substitutes of classical NAGS that are not OATs[14, 15]; These substitutes were able to complement *E. coli argA* mutants. They could be found in two forms in the genome of the organisms that contain them: fused to the ASL gene or as independent genes[15].

In some organisms like *Thermus thermophilus*, *Streptomyces coelicolor* or *Mycobacterium tuberculosis*, the lack of *argE* and *argA* and the possibility that their OATs are monofunctional probably makes them rely on S-NAGS for Arg synthesis[13]. S-NAGS has been characterized in *M. tuberculosis*[16]. This enzyme shows an extremely high K_m for glutamate (>600mM) which suggest that another enzyme could provide an efficient glutamate binding site, such as NAG kinase (NAGK), that would form an effective active acetylating complex as it appears to be the case in *Saccharomyces cerevisiae*[17], although in this organism NAGS is of the classical type except for the presence of a C-terminal extension of ~150 residues[18].

Acetylation of glutamate appears to control the *"in vivo"* flow of acetylated precursors that lead to Orn in those organisms that use AOD instead of OAT, and therefore that have a linear pathway. Classical NAGS activity is controlled through feedback inhibition by Arg[19], as well as is the case of S-NAGS[16]. The control of the activity of mono- and bifunctional OATs by arginine has not been fully proven, but Orn behaves as a competitive inhibitor[20] controlling the activity of these enzymes. Although significant progress has been achieved in the determination of the mechanism by which Arg inhibits classical NAGS and NAGK (see next section), such mechanism remains elusive for S-NAGSs.

1.5 Catalysts of the second or kinase step of the arginine biosynthetic pathway

The second step of the Arg biosynthetic pathway is the phosphorylation of NAG by the enzyme NAGK (Figure 2). In contrast with the first committed step of the pathway, this second step can only be catalyzed by NAGK, highlighting the importance of this enzyme and its regulation in the homeostatic control of the Arg pool in the organisms.

In some organisms NAGK is Arg-insensitive (*E. coli* and *Microbacterium lacticum*, among others[12, 21]), therefore NAGS becomes the most important point of control of the pathway, but in those microorganisms in which NAG is formed by transacetylation and therefore have a "cyclic" Arg biosynthetic pathway, NAGK is a crucial metabolic flow control point of the pathway[9].

One of the most frequent ways to control NAGK activity is through inhibition by the end product of the pathway Arg. The structural bases of such inhibition have been described by the group of Rubio[22] for the NAGKs from *Thermotoga maritima* and *P. aeruginosa*. Upon binding of Arg to the enzyme a conformational change occurs that stabilize NAGK in a more "open" state (T-state), in which the active site is enlarged consequently increasing the distance between the nucleotide and NAG, thus rendering a much less active NAGK (Figure 3).

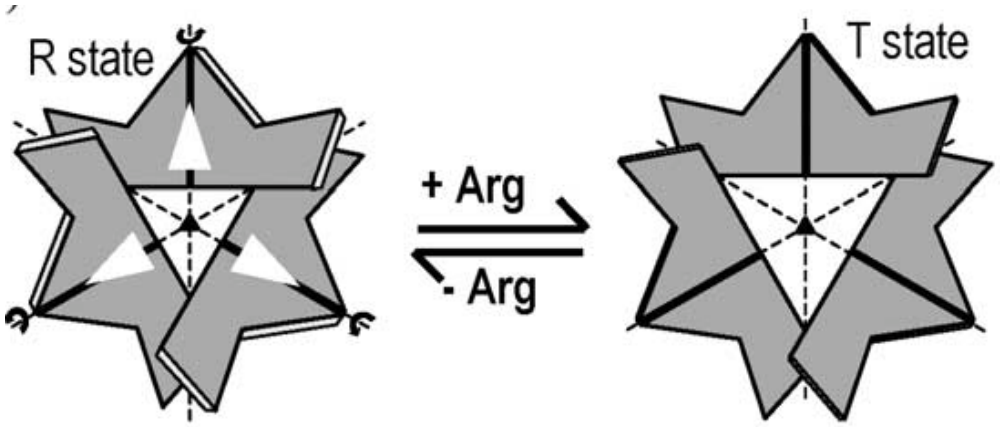


Figure 3: Conformational changes induced by the binding of Arg to hexameric NAGK. Adapted from Ramón-Maiques et al 2006[22]. Schematic representation of the R- to T-state transition induced by Arg upon binding to NAGK.

Since classical NAGS consists of a domain that is homologous to NAGK and a C-terminal smaller domain belonging to the GNAT family, the similarities of NAGK with the NAGK-like, or amino acid kinase (AAK) domain of NAGS led to the identification in classical NAGS of the Arg binding site, corresponding to that observed in NAGK, suggesting that the same inhibitory mechanism is operative for NAGK and NAGS. Recent studies from the same group involving site directed mutagenesis on *P. aeruginosa* NAGS, supported this hypothesis[23, 24].

In plants and photosynthetic microorganisms the control of NAGK activity is exerted not only by arginine, but also by the nitrogen abundance-signaling protein P_{II}[25]. In this case P_{II} binds NAGK reducing its sensitivity for Arg and allowing this amino acid to be synthesized continuously when nitrogen carbon and ATP are abundant, and to be stored in arginine-rich proteins, as in seeds, or in arginine-rich

polymers like cyanophycins, as in some cyanobacteria[26].

In ascomycetes (including, for example, yeast and *Neurospora crassa*) NAGS and NAGK are associated into a complex whose integrity has been proven to be essential for NAGS building as a stable protein, as well as for NAGS activity and effective inhibition by Arg[17].

1.6 The role of transcarbamylation in arginine synthesis and catabolism

Until now only the first two steps of the Arg biosynthetic pathway have been discussed. Although each and every step of the pathway is essential for its completion, transcarbamylation of Orn to produce Cit (Figure 2) represents yet another crucial step in the pathway (Figure 1).

The reaction catalyzed by OTC can participate in either the anabolism or the catabolism of Arg. While the carbamylation of Orn, yielding Cit, is involved in the biosynthetic sequence; the reverse reaction, the phosphorolysis of Cit, is the second step of the arginine deiminase (ADI) pathway (Figure 4). ADI, OTC and carbamate kinase (CK) convert Arg to Orn, ammonia and CO_2 in three steps producing ATP, which makes the ADI pathway a major source of energy for several organisms under anaerobic conditions[27-31].

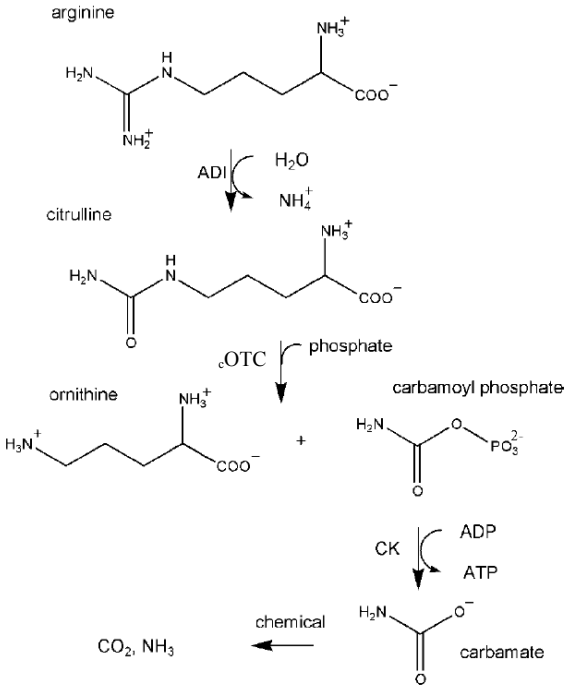


Figure 4: Schematic representation of the ADI pathway. Arg is converted to Orn, ammonia and CO_2 in three consecutive steps by ADI (arginine deiminase), cOTC (catabolic ornithine transcarbamylase) and CK (carbamate kinase), generating 1 mol of ATP per mol of Arg that enters the pathway.

Organisms that carry out both, arginine synthesis and ADI-pathway mediated arginine catabolism, generally have distinct anabolic and catabolic OTCases like it is the case of *Pseudomonads*[32-35]. Both OTCases (anabolic and catabolic) have been studied in detail in *P. aeruginosa*. In this organism anabolic OTCase, encoded by the gene *argF*, is repressed by Arg[36], while the catabolic enzyme, encoded by the *arcB* gene, which belongs to the *arcDABC* operon, is induced by Arg[37]. The two enzymes share ~52%[36] identity at the nucleotide sequence level. Despite their sequence similarities, the active form of ArgF is a homotrimer of approximately 100kDa while ArcB forms dodecameric tetramers of trimers of 456kDa[36, 38]. These differences in quaternary structure can apply also to other OTCases found in different organisms containing both anabolic and catabolic enzymes[33]. Catabolic OTC (cOTC) is supposed to be the limiting step in the ADI pathway[39], due to the irreversibility of the reaction catalyzed by ADI and the significant favor towards the formation of ATP of the CK reaction. In fact, of the enzymes of the ADI pathway of *P. aeruginosa* only catabolic OTC appears to be under regulation at the level of enzyme activity[40].

Anabolic OTC is widespread throughout phyla and is considered an essential arginine biosynthetic enzyme. Absence of the OTC gene or enzyme deficiency due to mutations causes Arg auxotrophy[5] in microorganisms and is lethal in ureotelic animals[41]. Among microorganisms, only a few obligate parasitic bacteria lack the OTC gene[42].

1.7 Transcarbamylation across phyla

The transcarbamylases family of enzymes belongs to the carboxyl- and carbamoyltransferases group, EC 2.1.3, a family that comprises enzymes that catalyze the transfer of a carbamoyl group from CP to an amino or oxygen group of a second substrate.

Members of the transcarbamylase family can be identified based on sequence identities in the N-terminal or CP-binding domain as all the members of this family share common residues involved in the binding of CP to the enzyme[43]. The most widely studied members of the transcarbamylases family are ornithine transcarbamylase (OTC) and aspartate transcarbamylase (ATC).

OTC and ATCase catalyze analogous reactions, and, as already indicated for OTC, are quite universally distributed[42]. While OTC utilizes Orn and CP, ATCase catalyzes the reaction between aspartate and CP that is the second step in the pathway of pyrimidine biosynthesis[44].

In 1999 Labedan and others[42], studied the evolutionary history of transcarbamylases using a limited number of sequences representing the three domains of life (44 OTCases and 33 ATCases). The results obtained from that work could not establish the actual order of organismal descent due to the complex topology of the composite rooted phylogenetic tree. The authors explained such complexity by assuming the occurrence of four subfamilies, two per group of enzymes; OTC α and OTC β and ATC I and ATC II, based on unique sequence signatures.

These studies on the evolution of the transcarbamylases indicated the

presence of such activity in the Last Universal Common Ancestor (LUCA), and ruled out the possibility of an evolutionary mechanism of fusion of independent modules. Despite the increasing number of sequences available, the robustness of the evolutionary tree constructed was confirmed by studies performed with ATC by the same laboratory[45] using quaternary structures to assess the evolutionary history of proteins. The general tree topology was not modified, simply the ATC I and ATC II groups were enlarged by the addition of new sequences but not dislodged[45].

The two groups of OTCases (OTC α and β) identified by Labedan et al.[42] were also confirmed by a more recent evolutionary study focused on OTC[46]. In their work, Naumoff et al.[46] highlighted the existence of sequences in the public databases, annotated as OTC that either did not catalyzed the transcarbamylation of Orn to form Cit (like the *Bacteroides fragilis* transcabamylase-like enzyme, previously crystallized by our group[47], therefore annotated as unknown transcarbamylase or UTC), or did catalyze the transcarbamylation of a substrate other than Orn (like the case of putrescine transcarbamylase, PTC). The evolutionary tree of transcarbamylases constructed by Naumoff et al.[46] using the 245 sequences for OTCases available at the moment of making the tree shown in Figure 5.

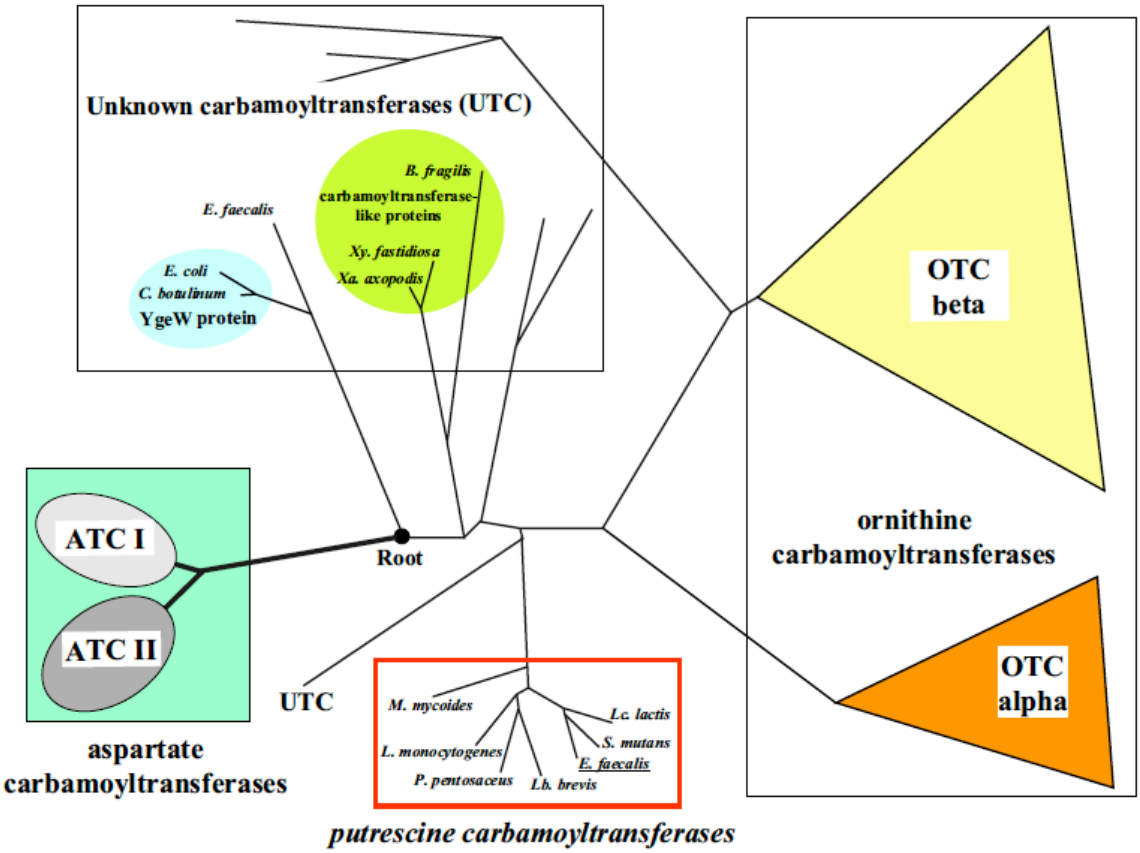


Figure 5: Evolutionary tree of transcarbamylases. Adapted from Naumoff et al. 2004[46]. A few paralogous ATCases belonging to either family ATC I or ATC II were used to root the tree. A black dot indicates the root. OTC families are schematized as triangles. The homologous transcarbamylases that are branching far from these ATC or OTC families are labeled as unknown (UTC). The few UTC sequences annotated as transcarbamylase-like are enclosed in ovals.

The functional assignment of proteins identified in genome sequencing projects is a major problem in post-genomic biology, as highlighted in the works of Naumoff et al.[46] and Sekowska et al.[48] among others[49-51]. Putative amino acid sequences have, in many cases, been annotated uniquely by a two-step process:

1. - Detection of a homologous relationship by a pairwise sequence similarity search at the level of primary structure
2. - Inference of functional similarity from this detected homology.

The work of Naumoff and others[46] was centered in the case of the PTC family of proteins that has been erroneously annotated as OTC and that form a distinctive branch from both ATC and OTC families (Figure 5). As Figure 5 shows, next to the branch formed by the YgeW protein from *E. coli* and *Clostridium botulinum* there is another branch that contains the sequence of a protein that has been reported to be essential for arginine biosynthesis in *Xanthomonads* and *Bacteroides*[47].

1.8 Sequence/structure traits and the substrate specificities of transcarbamylases

Transcarbamylases, as stated before, share a common functional domain, the CP-binding domain located in the N-terminal moiety[43] with significant proportions of identical or similar residues. The highest conservation across the transcarbamylase superfamily involves three regions: the SxRT, the HPxQ and the HxLP motifs[52]. All of them are involved in direct interactions with the common substrate for all the members of the family, CP. These homologies in the CP-binding domain have been used for the identification of the members of this family.

Furthermore comparison of the structures of OTC and ATC allowed identification in the CP-binding domains of both enzymes of residues involved in binding CP[53, 54]. In both active centers the STRT motif binds the phosphate oxygens of CP[55], and the active site also contains one R and one H residues (Arg^{106} and His^{136} in the human OCTase), which bind the carbonyl oxygen of CP[56]. In both families, at least one residue of the CP site is contributed by an adjacent subunit, in the case of *E. coli* ATCase, Ser 80 and Lys 84 of the adjacent subunit participate in the CP site, while in *E. coli* OTCase Gln 82 is the residue belonging to an adjacent subunit[56] (Figure 6).

Despite the structural similarities between the C-terminal moieties of OTC and ATC[57, 58], the sequence of this domain is more divergent than that of the N-domain, due to the variety of substrates that can be carbamylated[56]. Figure 6 illustrates the different interactions and residues involved in the binding of Orn and aspartate to OTC and ATC respectively.

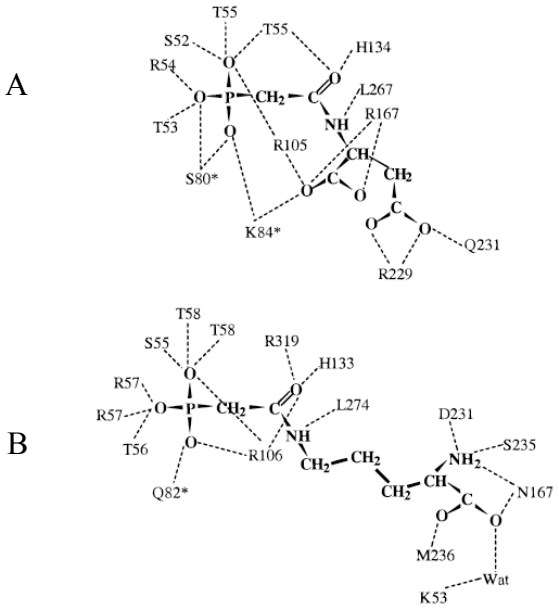


Figure 6: Schematic representation of the interactions between the substrates and the active site in *E. coli* ATC and OTC. Adapted from Allewell et al. 1999[56]. A) Interactions of the bisubstrate analog N-Phosphonoacetyl-L-Aspartate (PALA) with *E. coli* ATC. B) Interactions of the bisubstrate analog N-Phosphonoacetyl-L-Ornithine (PALO) with *E. coli* OTC. Residues from an adjacent subunit are labeled with an asterisk (*).

The reaction catalyzed by transcarbamylases involves the nucleophilic attack of the carbonyl C of CP by an amino group. Catalysis of this attack involves interactions mediated by the highly conserved HCLP motif of OTCases or the HPLP motif of ATCases. The difference between these two conserved motifs has to do with the charges of each of the substrates utilized by these enzymes[56]. This HCLP motif was thought to be a signature for Orn binding to OTC but the same motif is also present in PTCases[46].

Another sequence signature associated with transcarbamylases can be found in the so-called 240's loop (residues 236 to 259, *E. coli* OTC numbering) DxxxSMG in OTCases and RxQxER in ATCases. This loop is flexible and is an important ornithine or aspartate binding element[52]. The flexibility of this loop is essential for such binding since it requires a conformational change to create an induced-fit pocket in which the substrate is located.

The catalytic mechanism of both OTC and ATCases involves a tetrahedral intermediate. Glutamine 136 (Gln^{136}) in the case of *E. coli* OTC fulfills the role of stabilizing this intermediate, while Gln^{137} plays the same role in E. coli ATC. The region surrounding this glutamine in both OTC and ATC is strongly conserved; NxLxxxxHxxQxxD for OTCases and NxGDGxxxHxxQxxD in ATCases[56].

All the evolutionary conserved motifs present in the transcarbamylase family have led to the erroneous annotation of a large, yet increasing, number of genes as OTCases without experimental confirmation of such activity as has been mentioned before in this introduction and it has been highlighted by the studies of Labedan and collaborators[45, 46, 59]. The case of the transcarbamylase-like gene found in *B. fragilis* illustrates this statement.

1.9 Implication of a transcarbamylase-like gene with no OTC activity from *Bacteroides fragilis* in the *de novo* arginine biosynthesis in this organism.

While studying the arginine biosynthetic pathway in the anaerobe *B. fragilis*, using transposon-mediated mutagenesis Dr. Malamy, discovered the presence of a transcarbamylase-like (*BfargF*) gene essential for the biosynthesis of arginine in this organism (unpublished results).

In 2002, Shi et al.[47] described the crystal structure of this transcarbamylase-like protein from *B. fragilis* strain TM4000. The general fold of this protein (Figure 7) shared a high degree of structural similarities with anabolic OTCases and ATCases. However, there were several striking differences in some of the regions of the active site between the *Bf*ArgF' and OTC and ATC[47]

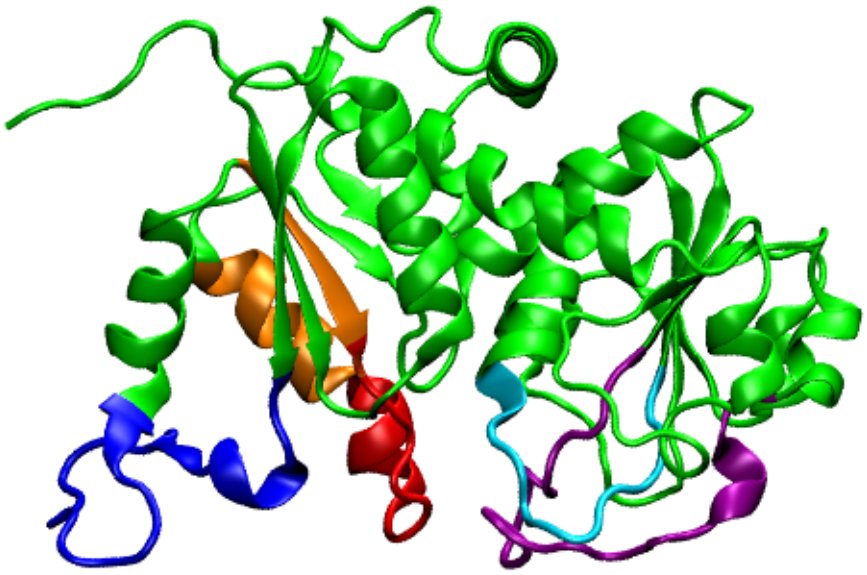


Figure 7: Ribbon diagram representation of the monomer structure of the transcarbamylase-like protein from *B. fragilis*. The 80's loop is shown in blue, the 120's loop in red, β-sheet B4 and α-helix H4 in orange, the 240's loop in purple and the proline-rich loop in cyan.

The 80's loop (residues 70-92, shown in blue in Figure 7) is 12 amino acids longer than that of OTCase or ATCase, and contains two short parallel strands (residues 77-78 and 91-92) that are unique to this protein. The 120's loop (residues 111-124, shown in red in Figure 7), which links B4 and H4 (shown in orange) near the active site, is ten residues longer than the respective loop in OTCase and ATCase. As a result, this loop extends further along the active site and is able to interact with the 240's loop (residues 236-259, shown in purple in Figure 7). The 240's loop, which is involved in recognizing the second substrate in OTCase and ATCase, and in triggering conformational changes upon binding of such substrate,

lacks the characteristic motifs DxxxSMG or RxQxER of OTCase and ATCase respectively, and has a different three-dimensional structure[47].

Even more surprising is the presence of a proline-rich loop (residues 173-183, shown in cyan in Figure 7) that links B6 and H6 and is located between the active site and the 240's loop[47]. Because of the number of prolines, this loop is not mobile, and would be expected to prevent movement of the 240's loop towards the active site when the second substrate binds. Although the structure of *Bf*ArgF' was solved, its enzymatic activity could not be detected.

According to the authors there were at least three possible ways in which this protein could be involved in producing citrulline that would explain why its activity has not been detected:

1. The enzyme requires an unidentified cofactor to bind the protein to become active.
2. The enzyme, within the cell, is part of a complex with other proteins and the disruption of this complex eliminates its activity.
3. The enzyme converts a substrate other than ornithine into an unknown intermediate that is subsequently converted to citrulline.

1.10 Identification of misannotated transcarbamylases

The fact that the transcarbamylase-like gene from *B. fragilis* was implicated in the *de novo* arginine biosynthesis in this organism, and the fact that the three-dimensional structure of the protein coded by *Bfargf'* was very similar to that of OTC and ATCases, increased the interest of our laboratory to study the transcarbamylase activity of this protein. Unfortunately initial attempts to elucidate such activity were unsuccessful[47].

Screening of the, at the time, available genomes using the *Bfargf'* sequence as a query allowed the identification of a similar gene in *Xylella fastidiosa*. Both genes shared a high degree of sequence similarity (52%) and homology (38%) and both genes lack the OTC and ATC signature sequences (DxxxSMG and RxQxER respectively)[47] (Figure 8). The predicted three-dimensional structure of the protein identified in *X. fastidiosa* was very similar to that of the *Bf*ArgF' indicating a similar function in the arginine biosynthetic pathway in this organism.

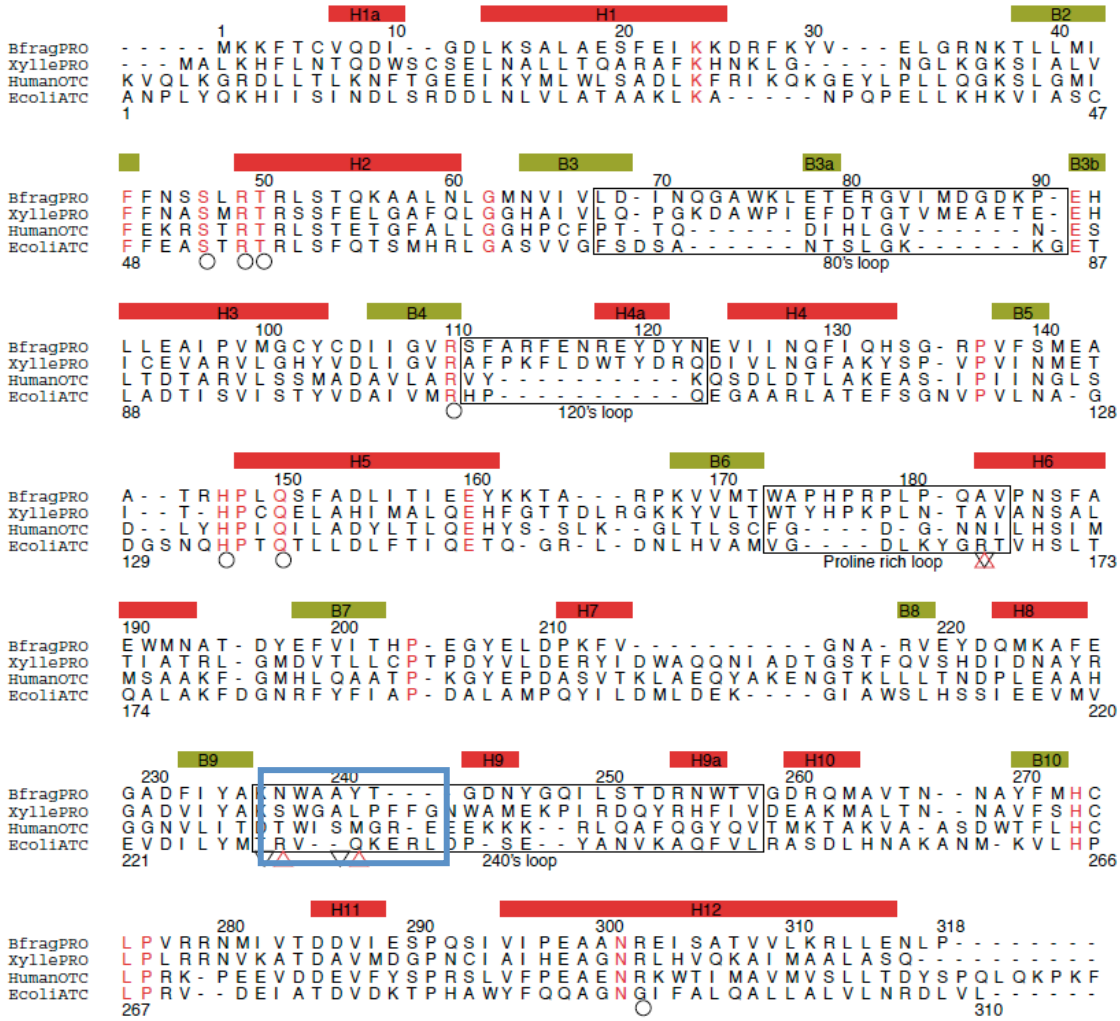


Figure 8: Amino acid sequence alignment of *B. fragilis* and *X. fastidiosa* transcarbamylase-like proteins and human OTC and *E. coli* ATC. Adapted from Shi et al. 2002. Structural elements are indicated by colored boxes above the sequence; green: β-sheets and red: α-helices. Red letters represents the residues conserved in all known transcarbamylases. Blue box highlights the region where the sequence signatures for both OTC and ATC can be found, which are not present in these other transcarbamylases of unknown activity.

Dr. Morizono carried out the initial bioinformatics analyses on the genes annotated as OTCases in the public protein database NCBI (www.ncbi.nlm.nih.gov/protein) to identify other members of this group of transcarbamylase-like proteins that lack the essential DxxxSMG motif.

A query using the keyword 2.1.3.3 (the E.C. number for OTCase) with NCBI Entrez returned 474 protein sequences (nowadays using the same keyword the search would return 3513 sequences of genes annotated as OTCases); those that appeared truncated or incomplete were discarded. The remaining genes were then screened for the presence or absence of the SMG motif using the EMBOSS module fuzzpro. Of the 474 sequences annotated as OTCase, 56 did not contain the SMG motif. These were aligned using pileup and the cluster associated with *B. fragilis* was used for further analysis. Figure 9 shows the sequence logo representation generated using WebLogo[60] (http://weblogo.berkeley.edu/), from the alignments of the C-terminal domain of those protein sequences that lacks the SMG motif

compared to the alignment of known OTC sequences. Among the *argF'* genes identified that lacked the SMG motif and annotated as potential transcarbamylases were those of *Bacteroides thetaiotaomicron*, *Cytophaga hutchinsonii*, several pathovars: *Xylella fastidiosa*, *Xanthomonas campestris*, *Xanthomonas axonopodis*, *Prevotella ruminicola*, and *Tannerella forsythus*[61].

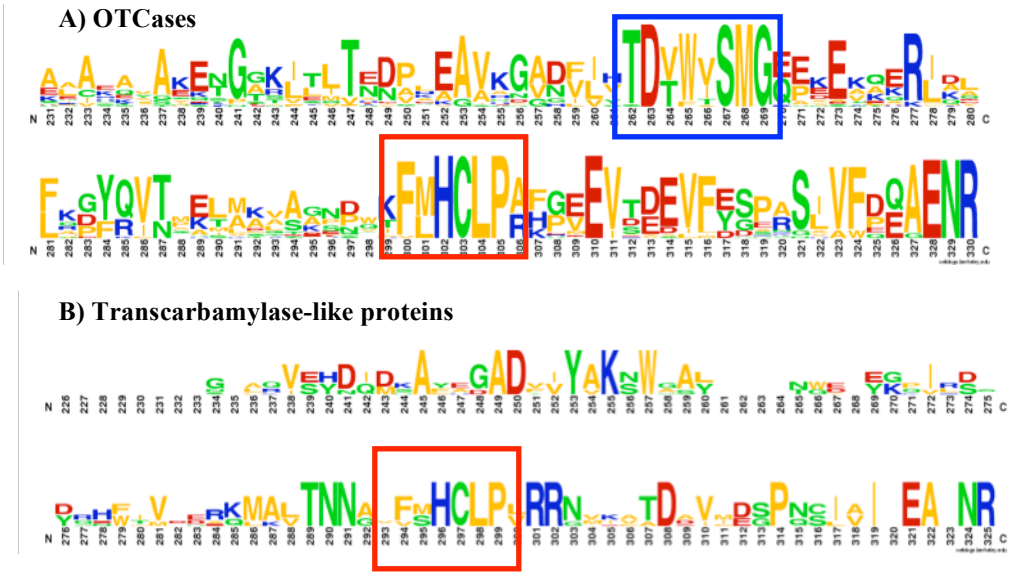


Figure 9: Representation of the alignment of known OTCases and transcarbamylase-like proteins identified in the initial bioinformatics analyses. In this figure the size of the letters is related to the degree of conservation among the sequences aligned and the color the type of residue. The motifs conserved in both groups are highlighted with red boxes. OTCase signature motif, present in all known OTCases and absent in the transcarbamylase-like proteins, is highlighted with a blue box.

These analyses showed the widespread distribution of these transcarbamylases of unknown activity among bacteria. In an effort to identify the activity of these transcarbamylase-like proteins, several *argF'* genes from these organisms were cloned, including: *Xylella fastidiosa*, *Cytophaga hutchinsonii*, *Xanthomonas campestris* and *B. fragilis argF'*.

Complementation studies carried out by Dr. Malamy, showed that *Xa. campestris*, *Xy. fastidiosa* and *C. hutchinsonii argF'* genes were able to revert the arginine auxotrophy of *E. coli* C6.1 (a K12 strain lacking both OTC genes *argF* and *argI*) when introduced by transformation into these cells and induced with IPTG. Thus, the ArgF' protein from these organisms functions in *E. coli* and provides a reaction product that can be used for arginine biosynthesis. However, when OTCase activity of these proteins was assayed with CP and ornithine as substrates no significant citrulline production was detected (data not shown).

The ability of all these proteins to complement *E. coli* C6.1 arginine auxotrophy suggested that the activity of these enzymes should be located upstream the reaction catalyzed by OTC in the pathway (Figure 2).

The first enzymatic activity discovered among these proteins was the N-acetylornithine transcarbamylase (AOTC) activity of *Xa. campestris* ArgF'. Which was elucidated using a coupled assay in which the lysate from *E. coli* C6.1 cells expressing $_{Xc}$ArgF' was incubated with each of the substrates of the enzymes CK, CPS, AOD and OTC

1.11 The specific context of this Thesis and of its objectives

The present thesis work is focused in the case of two families of misannotated microbial transcarbamylases that have been identified as essential for arginine biosynthesis, but which lacked the ability to catalyze the conversion of Orn into Cit.

The objectives of the present thesis work are the biochemical and molecular characterization of the AOTC activity of *Xc*ArgF' as well as the elucidation of the enzymatic activity of *Bf*ArgF' and its biochemical and molecular characterization.

We show that Xanthomonads and Bacteroidetes use novel pathways for the *de novo* arginine biosynthesis mediated by acetyl- and succinyl-ornithine transcarbamylases, we identify a recognition motif for these enzymes and we provide evidence supporting the view that the existence of succinyl-ornithine transcarbamylase requires that the first five steps of the arginine biosynthesis pathway use succinylated rather than acetylated intermediates.

Objectives

The objectives of the present PhD thesis are:

1. Biochemical characterization of ArgF' from *Xanthomonas campestris*.

- The specificity of the reaction will be investigated.
- The reaction products will be identified.
- The kinetics parameters will be determined
- The kinetic behavior of the enzyme in the presence of different specific inhibitors will be studied.

2. Design and synthesis of possible substrates as well as inhibitory compounds

- Possible substrates for the ArgF' enzyme from *B. fragilis* will be synthesized based on the data provided by the quaternary structure of the protein, as well as possible substrates for other enzymes implicated in the *de novo* arginine biosynthesis in these organisms.
- Inhibitory compounds will be synthesized based on the catalytic activity of the ArgF' enzymes from *X. campestris* and *B. fragilis*.

3. Identification and partial characterization of the activity of the enzyme codified by the *argF'* gene in *Bacteroides fragilis*.

- The activity of *Bacteroides fragilis* ArgF' will be identified.
- The activity of *Bacteroides fragilis* ArgF' will be partially characterized.

4. Determination of the implications of the presence of these transcarbamylases on the arginine biosynthetic pathway of the organisms that possess them.

- The logical metabolic flow of intermediates in the arginine biosynthetic pathway of *X. campestris* will be investigated.
- The initial steps of the arginine biosynthetic pathway from *B. fragilis*, leading to the intermediate metabolite that the ArgF' enzyme from this organism utilize will be investigated.

Some results described in this work have been published in the following peer-reviewed publications:

Morizono H, Cabrera-Luque J, Shi D, Gallegos R, Yamaguchi S, Yu X, Allewell NM, Malamy MH, Tuchman M. **(2006) Acetylornithine transcarbamylase: a novel enzyme in arginine biosynthesis. J Bacteriol. Apr; 188(8): 2974-82.**

Shi D, Morizono H, Cabrera-Luque J, Yu X, Roth L, Malamy MH, Allewell NM, Tuchman M. **(2006) Structure and catalytic mechanism of a novel N-succinyl-L-ornithine transcarbamylase in arginine biosynthesis of *Bacteroides fragilis*. J Biol Chem. Jul 21; 281(29): 20623-31.**

Shi D, Yu X, Cabrera-Luque J, Chen TY, Roth L, Morizono H, Allewell NM, Tuchman M. (2007) **A single mutation in the active site swaps the substrate specificity of N-acetyl-L-ornithine transcarbamylase and N-succinyl-L-ornithine transcarbamylase. Protein Sci. Aug; 16(8): 1689-99.**

Chapter 2: Functional characterization of a novel type of transcarbamylase that uses acetylornithine identified in *Xanthomonas campestris pv campestris*.

2.1 *Xanthomonas campestris pv campestris*

Xanthomonas is a genus that includes several plant pathogens. *X. campestris pv campestris* is a Gram-negative rod that attacks a variety of economically important crucifers, such as *Brassica* and *Arabidopsis*, causing black rot, which can appear on plants at any growth stage.

Disease control usually consists of using resistant cultivars, certified seeds, hot-water treatments of seeds followed by application of antibiotics of protecting fungicides, crop rotation, control of weeds and insects, and destruction of infected plants and debris in the field. Nowadays, the use of resistant cultivars is the most economical method of control[62].

X. campestris has also a commercial use since it is grown to produce the exopolysaccharide xanthan gum, which is used to control viscosity and as a stabilizing agent in many industries.

Phylogenetic analyses locate the *Xanthomonas* genus at the base of the g-subdivision of proteobacteria, within the same clade as *Xylella fastidiosa*. When its genome was sequenced only 64.91% of the genes were assigned a role category, the rest (35.01%) were either genes with not assigned role category (10.16%), conserved hypothetical genes (11.88%) and hypothetical genes (13.03%).

Essentially all the putative arginine biosynthesis genes in *X. campestris* and *X. axonopodis* can be identified by homology to known *arg* genes from various organisms; however, the sequences of their putative *argF* genes differ significantly from the canonical *argF* genes that encode OTCase enzymes[63]. These novel *argF'* genes are very similar to $_{BF}$*argF'* that is essential for arginine biosynthesis[47] as described in the Introduction.

As stated before in the Introduction the enzymatic activity of $_{Xc}$ArgF' has been identified to be that of N-acetylornithine transcarbamylase, a new member of the transcarbamylases family of proteins.

In this chapter I will present the results obtained for the Biochemical Characterization of the AOTCase activity of $_{Xc}$ArgF'.

2.2 Identification of the product of the transcarbamylase reaction catalyzed by *Xc*ArgF'

2.2.1 Preliminary identification using an amino acid analyzer of ACit as the product of the reaction catalyzed by *Xc*ArgF'

Previously described results indicated that the enzyme could use AOrn rather than ornithine and generates either ACit or citrulline. In the initial assay, *Xc*ArgF' activity was measured with a colorimetric method described by Pastra-Landis et al.[64] using known amounts of citrulline as standard.

This method detects the ureido group common to the side chains of both citrulline and ACit (Figure 10) and therefore could not discriminate between both of them, thus different approaches were necessary for the correct identification of the product of the reaction catalyzed by *Xc*ArgF'.

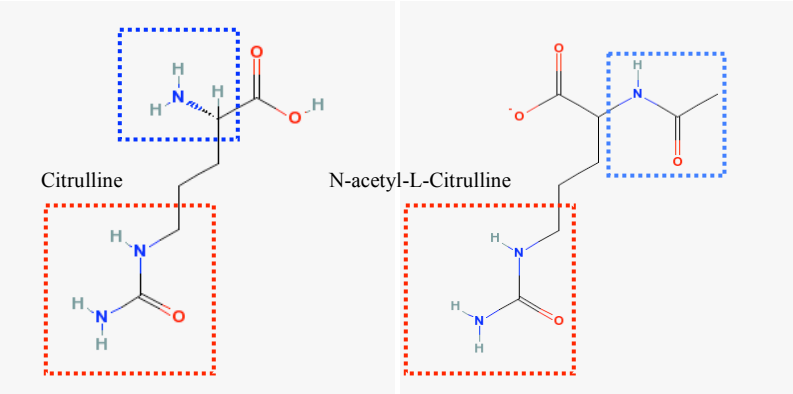


Figure 10: Schematic representation of the chemical structure of L-citrulline and N-acetyl-L-citrulline. Dashed red boxes show the ureido group and dashed blue boxes show the a-amino group in both compounds.

Ninhydrin is a compound that preferentially reacts with free amines and therefore, would react strongly with citrulline but weekly with ACit, which has a blocked α-amino group (Figure 10).

With the use of a Beckman amino acid analyzer (as described in Materials and Methods), the reaction of chemically and enzymatically synthesized ACit with ninhydrin was studied. Citrulline and AOrn have easily distinguishable different elution times (Figure 11), while ACit elutes near phosphoethanolamine and has a markedly reduced absorption coefficient at 570nm[65] (Figure 11), reflected on a lower absorbance at 570nm compared with similar amounts of citrulline.

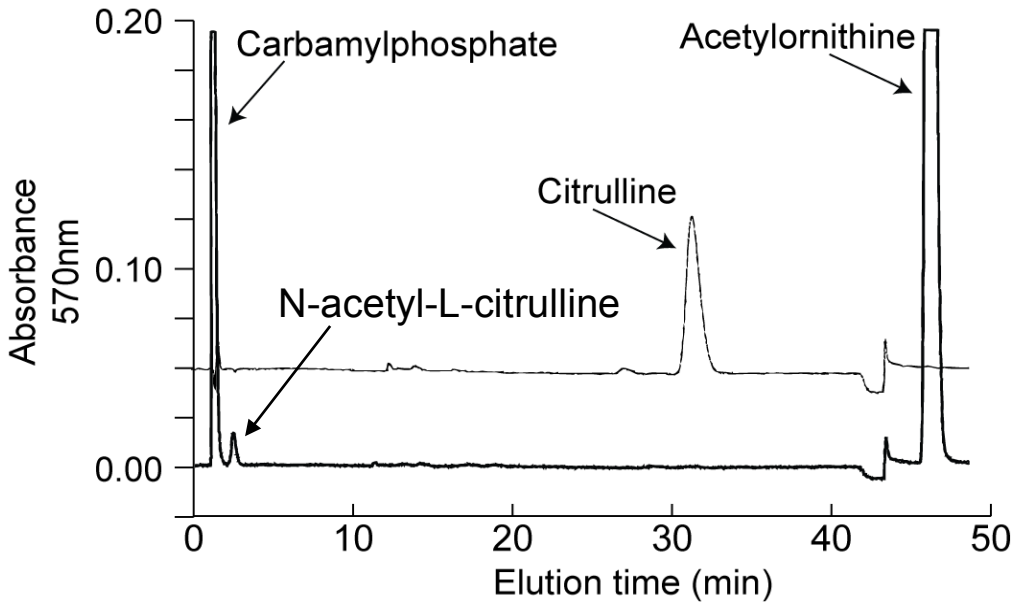


Figure 11: Representative amino acid analyzer chromatogram. For this assay 10µg of purified *Xc*ArgF' were incubated at 25°C in 1ml of 50mM Tris-HCl buffer pH 8.3 containing 5mM AOrn and 4.8mM CP. After 10 minutes 100µl of the reaction solution were diluted to 500µl with H2O and the reaction terminated with 500µl of 7% sulfosalicilic acid. 900µl of the color reagent (Materials and Methods: Colorimetric OTC activity assay) were added to the remaining 900µl of reaction solution for quantification of the product formed using the Colorimetric OTC activity assay. Based on the colorimetric assay, 136nmols of product were formed during the incubation. Thin line corresponds to the chromatogram generated by a solution of 110mM Cit (purchased from Sigma). The products of the reaction catalyzed by *Xc*ArgF' are represented with a thick line. For clarification of the figure thin lines are vertically offset for about 0.05 absorbance units.

If citrulline was the product of the reaction catalyzed by *Xc*ArgF' the chromatogram (thick line, Figure 11) should have shown a peak at the same retention time as the chemically synthesized citrulline (thin line, Figure 11), instead, it shows a peak with a retention time similar to that of phosphoethanolamine which is consistent with the expected results for ACit.

Although these results indicated that citrulline is not the product of the reaction catalyzed by *Xc*ArgF', they could only suggest that ACit is indeed the product formed from AOrn and CP in the presence of *Xc*ArgF' since no information about the chemical nature of the product of the reaction could be extracted from them.

We used liquid chromatography coupled with mass spectrometry (LC-MS) and nuclear magnetic resonance (NMR) to gather information about the chemical structure of the product of the reaction catalyzed by *Xc*ArgF'.

2.2.2 LC-MS and NMR analysis confirmed that the product of the reaction catalyzed by *Xc*ArgF' is N-acetyl-L-citrulline

Using the results of the colorimetric detection method described by Pastra-Landis et al.[64] to measure OTC activity as an estimate approximately equal amounts of both enzymatically and chemically synthesized ACit (commercially available) were analyzed by LC-MS (Materials and Methods).

Analysis of these samples revealed the presence of a peak with a mass equal to that of ACit in a protonated state (218.1 *m/z*) in both samples (Figure 12), citrulline in the same state would have been detected as an ion of 176.1 *m/z* and a different retention time. Boiled enzyme used as a control did not show this peak, neither the control without the enzyme (Figure 12).

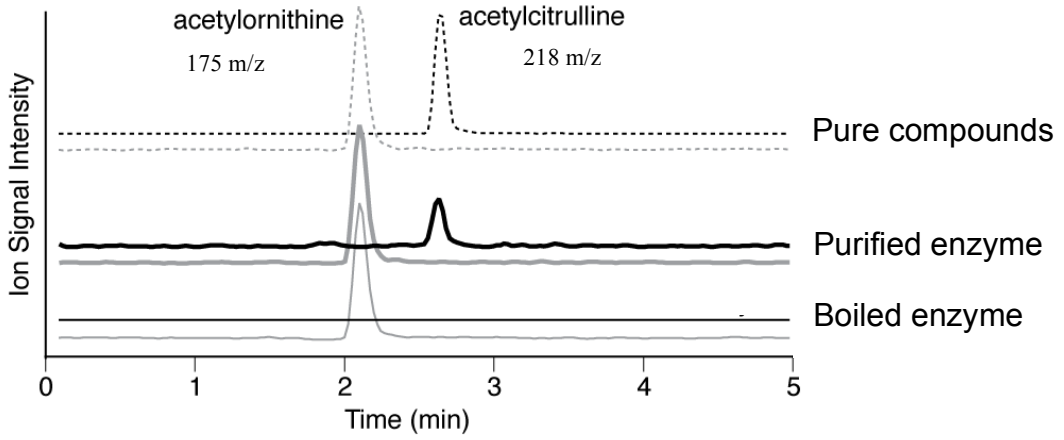


Figure 12: LC-MS analysis of the products of the reaction catalyzed by *Xc*ArgF'. For these analyses 2µg of purified *Xc*ArgF' were incubated for 5 minutes at 25°C in 1ml of 50mM Tris-HCl containing 1mM CP and 5mM AOrn. After the incubation time, 1ml of 30% tri-chloroacetic acid (TCA) was added to stop the reaction and precipitate the protein. Samples were then centrifuged and the supernatant was loaded onto the instrument without further purification (Materials and Methods). The chromatograms represented in the figure are those of the single ions corresponding to each of the compounds extracted from the total ions detected by the instrument during the run (Materials and Methods). Dashed lines: pure compounds (AOrn and ACit). Thick lines: enzymatically synthesized products by *Xc*ArgF'.

These results showed the presence of a compound with the same m/z and retention time than those of purified ACit among the products of the reaction catalyzed by *Xc*ArgF', which in conjunction with the previous results from the colorimetric assay and ninhydrin reactivity suggested that ACit is most likely to be the product of the reaction.

Final confirmation about the nature of the product of the reaction catalyzed by *Xc*ArgF' was obtained by NMR analyses (Materials and Methods).

Figure 13 shows the results obtained for the COSY-NMR (top) and two-dimensional COSY-NMR (bottom) spectra of both, chemically and enzymatically synthesized ACit.

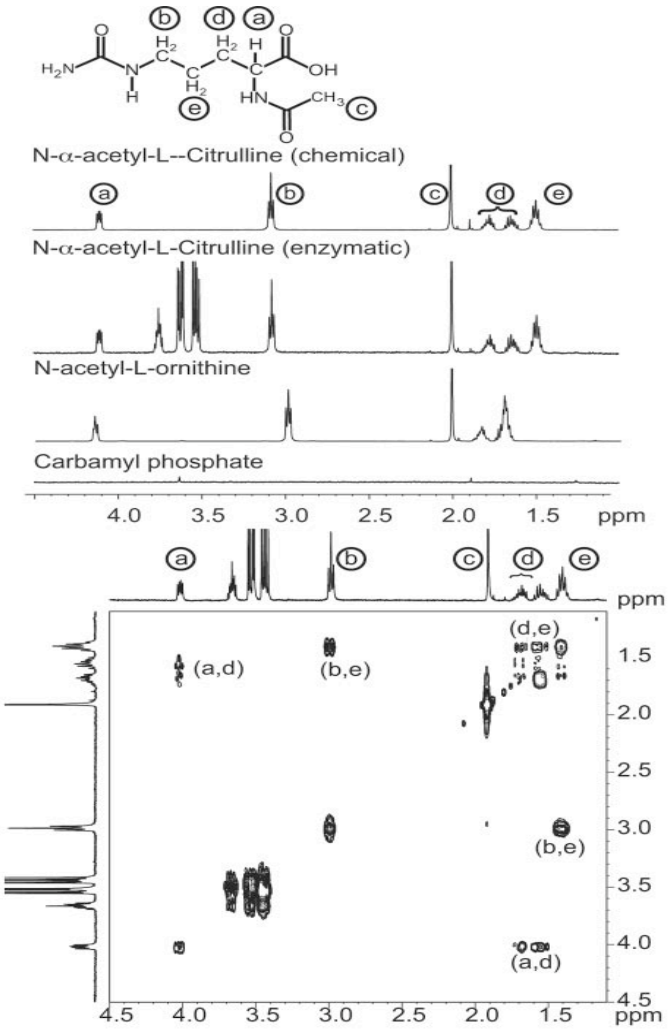


Figure 13: NMR analyses of the products synthesized by *Xc*ArgF' compared to the chemically synthesized compounds. COSY-NMR spectra show the enzymatically generated compound to be *N*-acetylcitrulline. (Top) The chemical structure of ACit with the protons seen in the labeled NMR spectra. 1H-NMR comparison of chemically and enzymatically synthesized ACit, AOrn, and CP. (Bottom) Two-dimensional COSY-NMR of the enzymatically synthesized ACit sample. The self-correlation peaks can be seen on the diagonal from lower left to upper right. Coupled peaks are seen off the diagonal. Protons that are coupled are identified by tracing horizontally and vertically off a point on the diagonal to a coupled peak and then returning vertically or horizontally to a new peak on the diagonal. The peaks between 3.5 and 3.8 ppm are not coupled to any other groups in *N*-acetylcitrulline and are likely due to a contaminant from the enzyme preparation

Both NMR analyses identified the product of the reaction catalyzed by *Xc*ArgF' as N-acetyl-L-citrulline thus supporting the previous results obtained by the different techniques presented in this section.

The data presented in this section provide evidence that the activity of *Xc*ArgF' is indeed that of a N-acetyl-L-ornithine transcarbamylase (AOTCase) and that the product of such activity is ACit. The reaction catalyzed by XcArgF' would be described by the following equation:

carbamoyl phosphate + N-acetyl-L-ornithine $\xrightarrow{{}_{Xc}\text{ArgF'}}$ **N-acetyl-L-citrulline + Pi**

Therefore, herein it will be designated as $_{Xc}$AOTCase.

2.3 Characterization of the AOTCase activity of $_{Xc}$ArgF'

2.3.1 Purification of ArgF' from *X. campestris pv campestris* ($_{Xc}$AOTCase)

$_{Xc}$ArgF' was purified as described by Shi et al. [63] using an AKTA FPLC by nickel affinity. This enzyme was purified to homogeneity with >96.5% purity.

Figure 14 shows a representative SDS-PAGE of the proteins presents in the fractions collected during the purification of this enzyme.

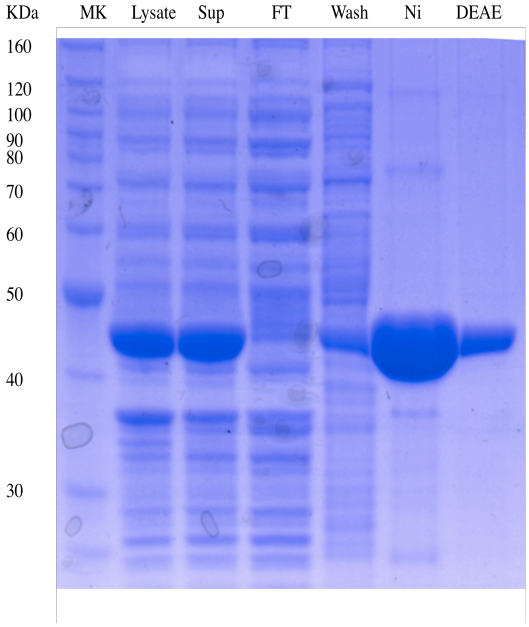


Figure 14: Representative gel showing the different stages of the purification of $_{Xc}$AOTCase. MK: Size marker (BenchMark Ladder), Sup: Supernatant (after cell lysis and membranes precipitation), FT: Flow through (unbounded proteins), Wash: weakly bounded proteins eluted with mild stringent buffer, Ni: fractions collected after elution of the protein bonded to the nickel column, DEAE: fractions collected after elution of the protein bind to DEAE column.

The molecular weight determined experimentally differs from that calculated from the 339 amino acids polypeptide encoded by the $_{Xc}$*argF'* gene (37.8 KDa). Such difference can be explained by the addition of a poly-histidine tag when cloning the gene.

After purification, all fractions containing the purified $_{Xc}$AOTCase protein were pooled and dialyzed in storage buffer (Materials and Methods) prior to its use for enzymatic assays.

Purified $_{Xc}$AOTCase was quantified using the Bradford method as described in Materials and Methods.

2.3.2 Specificity of the reaction catalyzed by $_{Xc}$AOTCase

To determine the specificity of the reaction catalyzed by $_{Xc}$AOTCase as well as to elucidate if the AOTCase activity could be assigned to any of the other transcarbamylases identified in this study, ornithine and AOrn were used as substrates for the transcarbamylase assay.

In this experiment, 2µg of purified $_{Xc}$AOTCase were assayed along with 2µg of purified ArgF' proteins from *C. hutchinsonii* ($_{Ch}$ArgF), *B. fragilis* ($_{Bf}$ArgF) and human OTCase ($_{h}$OTCase), using the standard conditions for OTCase assay[41], that is: 5 mM ornithine or AOrn and 4.8 mM CP (Figure 15). Boiled enzymes were used as negative controls and the product formation was quantified using known amounts of citrulline as standard.

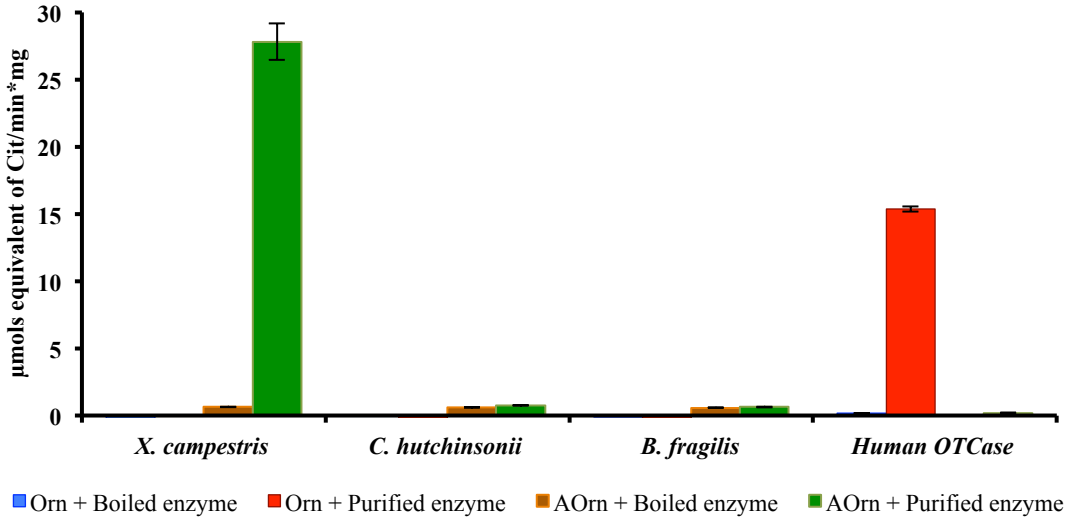


Figure 15: Comparison of the ability of different ArgF' enzymes to utilize ornithine or AOrn as substrates. 2µg of purified ArgF' from *X. campestris*, *C. hutchinsonii*, *B. fragilis* and *H. sapiens* were assayed in 1ml of 50mM Tris-HCl pH 8.3 in the presence of 5mM of ornithine or AOrn and 4.8mM CP. Incubations were carried out at 25°C in a water bath for 5 minutes. After 5 minutes 1ml of fresh color reagent (Materials and Methods) was added to stop the reaction. Tubes containing the samples were incubated over night at room temperature with constant shaking at 125rpm in the dark. The following day tubes containing the samples assayed were incubated at 45°C for 25 minutes for final color development. Sample absorbance was then measured spectrophotometrically at 466nm.

The results shown in Figure 15 suggest that the reaction catalyzed by $_{Xc}$AOTCase is highly specific for AOrn, and that it must play a crucial role in the arginine biosynthetic pathways in this organism since no other transcabamylase-like gene can be found in the genome of *X. campestris*.

In addition, these results prove that the AOTCase activity might be restricted to members of the Xanthomonads family of microorganisms and is not broadly distributed among the phyla identified in the initial bioinformatics analyses.

2.3.3 AOTCase assay conditions optimization

Salt concentration

High concentrations of salt (KCl, 200mM) has been proven to stabilize *Pirococcus furiosus* OTCase by preventing the dissociation, at high temperatures, of the dodecameric active estate of the protein[66], thus the effects that several concentrations of NaCl and KCl could have on the activity of *Xc*AOTCase were tested.

As shown in Figure 16, both salts, even at concentrations of 200 mM, have very little or no effect on the activity of *Xc*AOTCase under the experimental conditions tested (Materials and Methods).

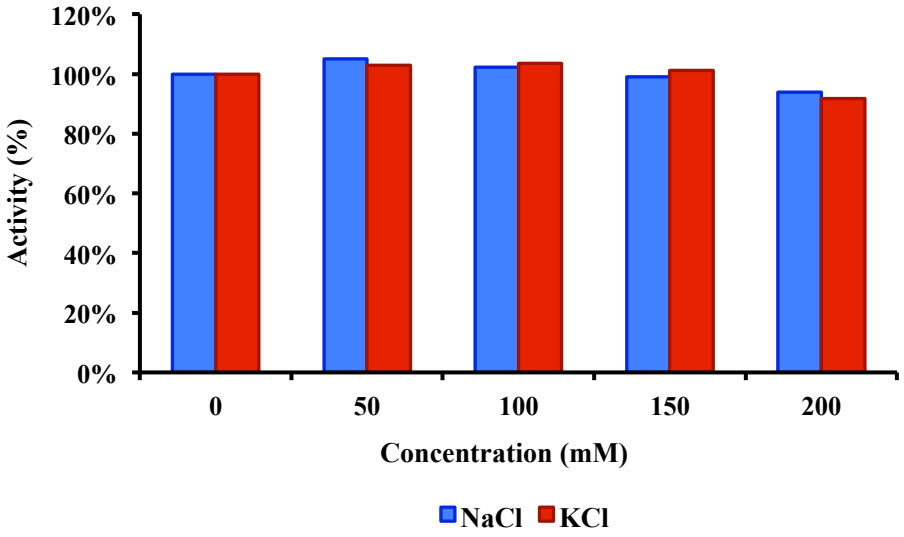


Figure 16: Effects of various concentrations of NaCl and KCl on *Xc*AOTCase activity. 2µg of purified *Xc*AOTCase were assayed in the colorimetric assay, in 1ml of 50mM Tris-HCl pH 8.3 containing 1mM AOrn, 4.8mM CP and either NaCl or KCl at different concentrations, at 25°C for 5 minutes. Percentages of activity were calculated using the activity measured in the absence of each salt, as reference.

These results show that the activity of *Xc*AOTCase is not dependent on the presence or absence of neither NaCl nor KCl, therefore, based on these results no salts were added to the experimental conditions to assay the activity of *Xc*AOTCase.

Enzyme concentration and time of reaction; Detecting enzyme inactivation

Determination of the optimal enzyme concentration as well as time of reaction are crucial in order to obtain reliable and reproducible results since certain enzymes that are stable under optimized conditions of long-term storage will inactivate during the course of an activity assay.

This behavior is characterized by progress curves that plateau early, before significant substrate loss has occurred. Enzyme inactivation commonly occurs for two reasons:

- the active conformation of the enzyme may not be stable under the specific conditions used in the assay and
- spontaneous enzyme inactivation that results from catalytic turnover. For some enzymes, the chemistry associated with turnover can lead to inactivation of the enzyme by covalent adduct formation, of by destruction of a key active site amino acid residue or cofactor.

Enzyme inactivation during the course of the time of reaction can be detected using the Selwyn test[67]. To perform this test, the reaction progress has to be measured at several different concentrations of enzyme. If the enzyme is not inactivated during the course of the reaction, plots of *p* (concentration of product after time) against e_0t (e_0: total enzyme concentration, t: time) should be superimposable

We performed the Selwyn test for *Xc*AOTCase as described in Materials and Methods and the results obtained are shown in Figure 17.

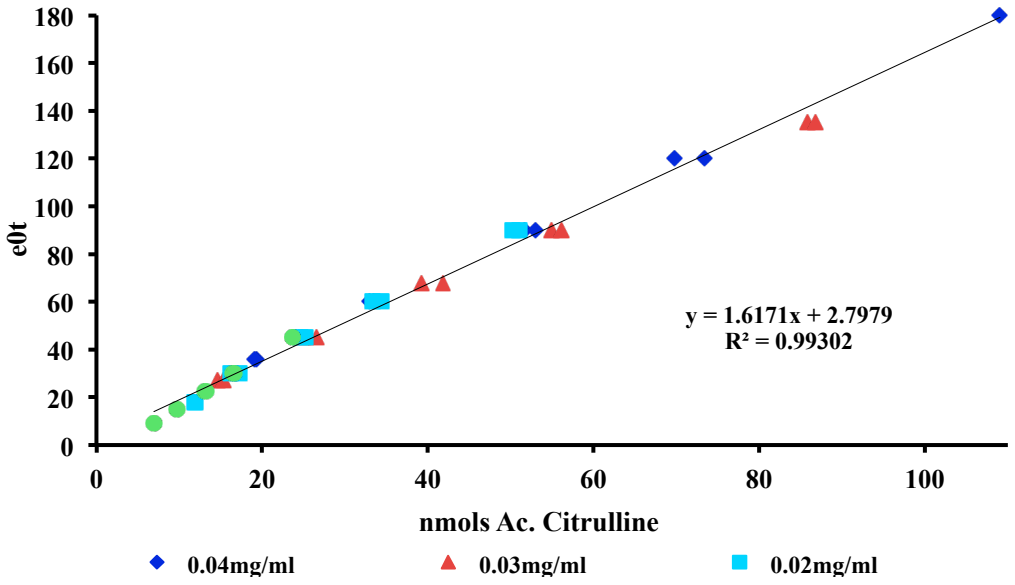


Figure 17: Detecting enzyme inactivation during the course of the reaction catalyzed by *Xc*AOTCase. Different amounts of purified *Xc*AOTCase were assayed as described before in 1ml of 50mM Tris-HCl pH 8.3 in the presence of 5mM AOrn at 25°C. The reaction was stopped by addition of fresh color reagent (Materials and Methods) at 3, 5, 7.5, 10 and 15 minutes after addition of 0.5mM CP to the reaction mixture. Linear regression fitting was performed using all the data obtained in this experiment.

These results demonstrate that the rate of product formation is linear with respect to time and independent of the amount of enzyme used in the assay.

2.3.4 Determination of the kinetic parameters that characterize the activity of *Xc*AOTCase

To characterize the reaction catalyzed by *Xc*AOTCase, the Michaelis constant (K_m) and the maximum velocity (V_{max}) were determined using saturation curves for both substrates (AOrn and CP) as described in Materials and Methods.

AOrn saturation curve

The AOrn saturation curve for *Xc*AOTCase was performed in the presence of 4.8mM of CP while varying the AOrn concentration from 0 to 5mM as Figure 18 shows. The reaction was carried out for 5 minutes at 25°C.

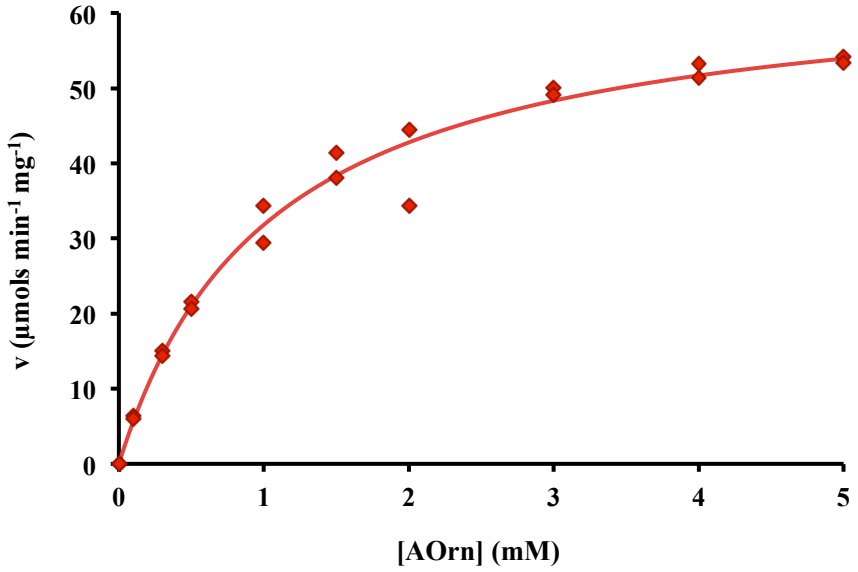


Figure 18: Determination of the kinetics parameters K_m and V_{max} for AOrn of *Xc*AOTCase. 0.3µg of purified *Xc*AOTCase were incubated as described before, in the presence of a constant amount of 4.8mM CP while varying the AOrn concentration. After 5 minutes the reaction was stopped and the product formation was quantified using ACit as standard. Data were fitted to the Michaelis-Menten equation for enzyme kinetics and the values determined for K_m and V_{max} were 1.05±0.23 mM and 65.08±4.77 µmols/min/µg respectively.

The reaction catalyzed by *Xc*AOTCase follows a Michaelis-Menten type of kinetics for AOrn (Figure 18). K_m and V_{max} were: 1.05 ± 0.23 mM and 65.28 ± 4.77 µmols $min^{-1}µg^{-1}$ respectively, as determined by non-linear fitting of the experimental data to the following equation:

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

using the program GNUPLOT (by Williams & Kelley, 2004).

These values are on the range of those obtained for other members of the OTCase family. In general, OTCases from other organisms have a lower affinity for ornithine than *Xc*AOTCase has for AOrn, with the exception of the human and *S. cerevisiae* enzymes in these cases the affinity for ornithine is 10 and 2 times greater respectively[41, 68]).

CP saturation curve

K_m for AOrn (1.05mM) calculated as described in the previous section was used as a reference for the determination of the kinetic parameters for CP. In this case the AOrn concentration was kept constant at 5mM (about 5 times the K_m), while the CP concentration was varied from 0 to 1mM.

As determined for AOrn, the reaction follows a Michaelis-Menten type of kinetic (Figure 19) although in this case the affinity for CP of *Xc*AOTCase is much higher than that for AOrn.

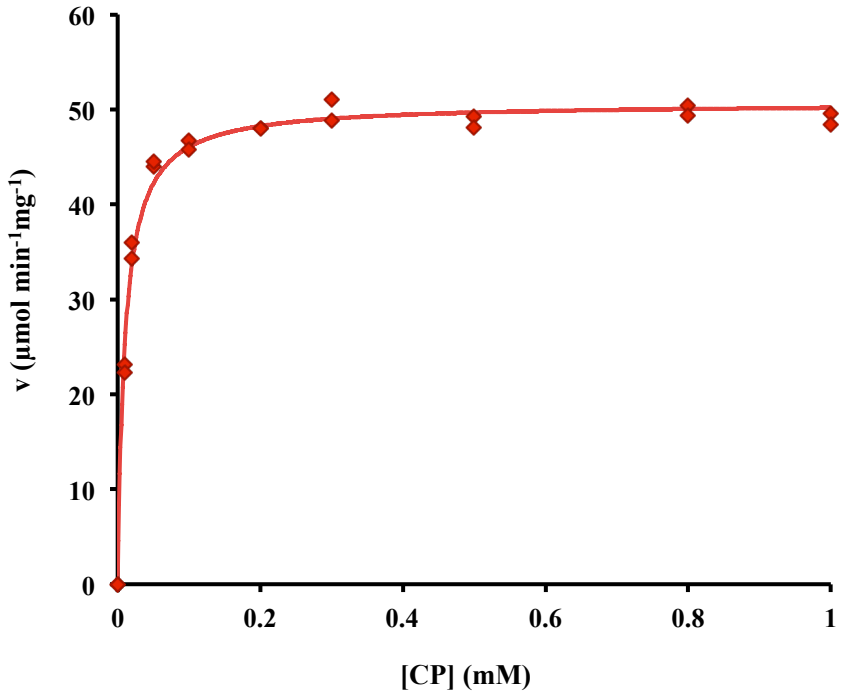


Figure 19: Determination of the kinetic parameters K_m and V_{max} for CP of *Xc*AOTCase. 0.3µg of purified *Xc*AOTCase were incubated as described before, in the presence of a constant amount of 5mM AOrn while varying the CP concentration. After 5 minutes the reaction was stopped and the product formation was quantified using ACit as standard. Data were fitted to the Michaelis-Menten equation for enzyme kinetics and the values determined for K_m and V_{max} were 0.01±0.0007mM and 50.68±0.48 µmols/min/µg respectively.

The kinetic parameters for CP were calculated as described in the previous section using the GNUPLOT software. The K_m value for CP of *Xc*AOTCase is 0.01mM, such affinity is the greatest compared to that of other members of the OTCases family for the same substrate, only *S. cerevisiae* OTCase possess an equivalent value (0.019 mM[68]). *E. coli* and human OTCases have an affinity 20 and 5 times lower[41, 69], respectively, while the difference with *P. aeruginosa* and *N. crassa* enzymes is 100 and 250 times lower respectively[70, 71].

The lack of activity of *Xc*AOTCase using ornithine as a substrate (Figure 15) suggest that acetylation of the aN-group must be necessary not only for the substrate of the reaction catalyzed by this enzyme but, as well, for potential inhibitors to be able to exert any kind of effect over *Xc*AOTCase.

The following section covers the study of the inhibitory effects of the acetylated forms of some of the most effective inhibitors of OTCases.

2.4 Inhibition studies on *Xc*AOTCase

2.4.1 Substrate inhibition

*Xc*AOTCase has shown a kinetic behavior analogous to other previously characterized OTCases. One of the characteristics common to several OTCases as well as ATCases is the inhibition by the reaction substrates ornithine or aspartate respectively[41, 72-74]. Sometimes like in the case of human OTC the inhibition is very pronounced even at concentrations above 1 mM of ornithine[41].

Since *Xc*AOTCase has a K_m for AOrn similar to that of other OTCases for ornithine we expected the same kinetic behavior thus we assayed the AOTCase activity of *Xc*ArgF' in the presence of high concentrations of AOrn while maintaining the CP concentration constant 10 times above the determined K_m (0.01mM) (Figure 20).

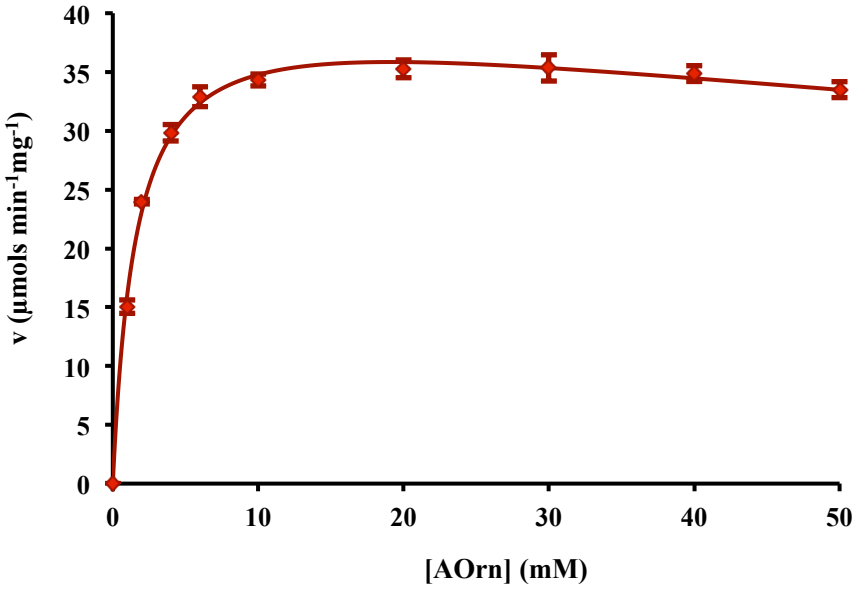


Figure 20: Studies on the inhibition of *Xc*AOTCase by the reaction substrate AOrn. The effects of high concentrations of AOrn on the activity of *Xc*AOTCase were studied using 2µg of purified enzyme in the presence of a constant concentration of 0.1mM CP while varying the AOrn concentration from 0 to 50mM. Assays were carried out at 25°C for 5 minutes in 50mM Tris-HCl pH 8.3 and the product formation was quantified using ACit as standard. Every assay was done in triplicate and each point represents the average and the standard deviation obtained for each condition.

Although not very pronounced, *Xc*AOTCase shows a slight deviation from Michaelis-Menten kinetics at high AOrn concentrations (up to 50 mM) (Figure 20). The results shown in Figure 20 were fitted to a model of substrate inhibition described by the following equation:

$$v = \frac{V_{\max}[S]}{K_m + [S] + \frac{[S]^2}{K_{si}}}$$

Where V_{max} is the maximum enzyme velocity if the substrate didn't inhibit enzyme activity, K_m is the Michaelis-Menten constant and K_{si} is the dissociation constant for substrate binding and represent the substrate concentration at which the velocity of the reaction is half the V_{max}.

The dissociation constant calculated for AOrn in the case of *Xc*AOTCase is 231mM, which is about 10 times higher than the equivalent value for ornithine inhibition of human OTCase[41]. These results emphasize the differences between this new member of the transcarbamylases family and the well known OTC and ATCases, highlighting the discrepancies between nucleotide sequence homologies and enzyme activity.

2.4.2 Substrates analogs inhibition studies

It has been shown that ornithine analogs such as L-norvaline (Norv) or L-serine (Ser) can inhibit the reaction catalyzed by *E. coli* OTCase[69], therefore structural analogs of AOrn are expected to behave as inhibitors of *Xc*AOTCase.

In this section I will present the results obtained from the study on the effects of N-acetyl-L-norvaline (ANorv) and pyrophosphate (PP), compounds that are structural analogs of AOrn and CP respectively.

N-acetyl-L-norvaline effects on XcAOTCase activity

Norvaline is an structural analog of ornithine that behaves as a competitive inhibitor of OTCase with a K_i of 5.8 µM at pH 7.9[72]. Due to the high specificity for AOrn of *Xc*AOTCase, norvaline was expected to have very little, if any, effect on its activity.

Inhibition studies in the presence of 5mM norvaline confirmed this hypothesis. This compound reduced the activity of *Xc*AOTCase only by 13% (Figure 21).

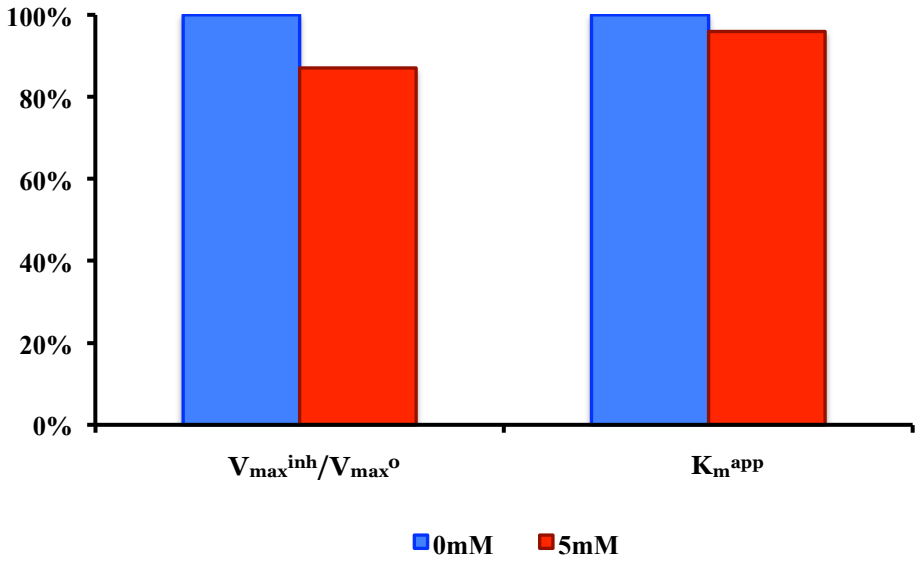


Figure 21: Inhibitory effects of norvaline on *Xc*AOTCase activity. *Xc*AOTCase activity was assayed in the presence of 0 or 5mM (blue and red bars respectively). Data represent the % of reduction of the V_{max} and K_m.

Therefore, the acetylated derivative of norvaline, ANorv, would be expected to inhibit *Xc*AOTCase in a similar way as norvaline inhibits OTCase.

To confirm this hypothesis, 2µg of purified *Xc*AOTCase were assayed as described in Materials and Methods using AOrn saturation curves, in the presence of different concentrations of ANorv, ranging from 0.125 to 1mM (Figure 22)

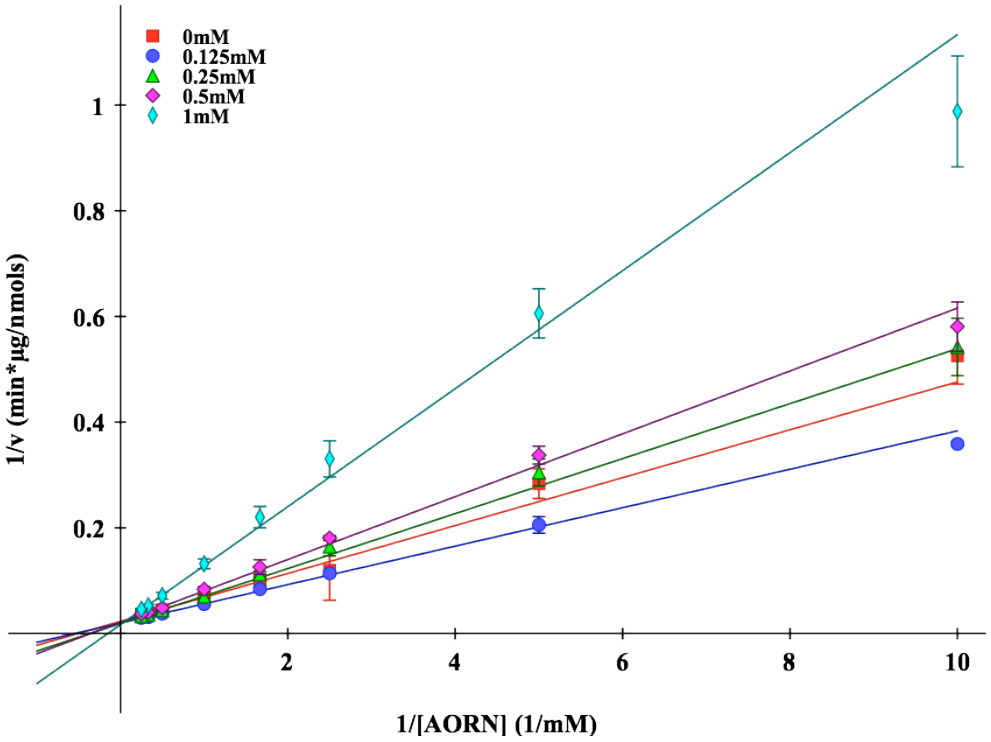


Figure 22: Studies on the inhibition of *Xc*AOTCase activity by the structural analog ANorv. 2µg of *Xc*AOTCase were assayed in 1ml of 50mM Tris-HCl pH 8.3 with a constant concentration of 0.1mM CP while varying the concentration of AOrn in the presence of different concentrations of ANorv ranging from 0 to 1mM. Lineweaver-Burk representation was chosen to more clearly illustrate the kinetic behavior. The trend lines were produced using the apparent V_{max} and K_m values derived from direct nonlinear least squares regression of the data. Each point shows the average and standard deviation of three independent assays.

The results obtained in this experiment showed that ANorv behaves as a competitive inhibitor of *Xc*AOTCase as we hypothesized previously.

The calculated K_i for ANorv with respect to AOrn is 0.3 mM, a value much higher than that of norvaline against *E. coli* OTCase. This difference can be explained by different interactions of the inhibitors to the respective enzyme.

Pyrophosphate effects on Xc AOTCase activity

CP is the common substrate to all the members of the transcarbamylases family, therefore compounds that attach to the CP binding site of the proteins are expected to affect the enzymatic activity of these proteins. Pyrophosphate (PP) a structural analogue of CP, is able to competitively inhibit ATCase and OTCases activity and thus it was expected to inhibit *Xc*AOTCase as well.

*Xc*AOTCase activity was assayed using 2µg of purified enzyme in the presence of different amounts of PP (0, 0.63, 1.25, 2.5 and 5mM) as described in Materials and Methods, maintaining constant the CP concentration while varying the AOrn concentration.

Figure 23 shows the effects of PP on *Xc*AOTCase activity expressed in variation of the kinetic parameters V_{max} and K_m.

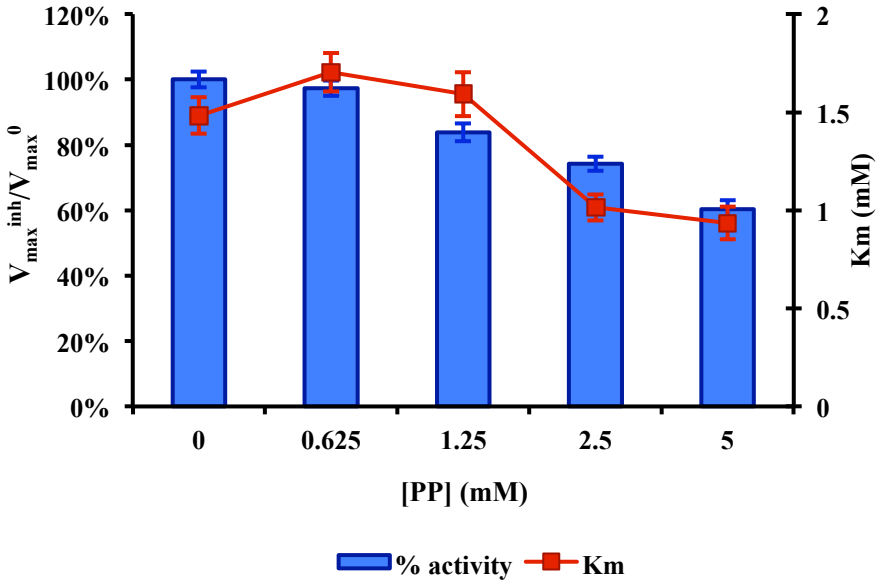


Figure 23: Effects of PP over *Xc*AOTCase activity. Data represent the % of reduction of the V_{max} (Blue bars, primary vertical axe) in the presence of 0.63, 1.25, 2.5 and 5mM PP compared to the V_{max} in the absence of inhibitor, and the effects exerted on the K_m at the same concentrations of PP (Red squares, secondary vertical axe).

Because of the results obtained under these conditions were not conclusive; the inhibitory mechanism of PP towards *Xc*AOTCase could not be determined.

2.4.3 Synthesis of acetylated PALO derivative α-N-acetyl-δ-phosphonoacetyl-L-ornithine (PALAO) and study of its effects on *Xc*AOTCase activity.

δ-N-phosphonoacetyl-L-ornithine (PALO) is an stable transition-state analogue (for CP and ornithine) and a potent inhibitor of OTCase (the K_i for *E. coli* OTCase is 0.77 μM[75]). As in the case of norvaline and ANorv, we hypothesized that the acetylated-PALO derivative, α-N-acetyl-δ-N-phosphonoacetyl-L-ornithine (PALAO) would effectively inhibit the activity of *Xc*AOTCase while PALO should have no effects.

Synthesis of PALAO

To test our hypothesis, we needed to synthesize PALAO from PALO.

As described in Materials and Methods, 5.1mg (0.02mmols) of PALO were dissolved in 200μl of water and incubated with 0.2mmols of acetic anhydride for 2

hours at room temperature with constant stirring. After incubation the solution was neutralized with 2ml of 1% (w/v) KOH.

Chemically synthesized PALAO was tested for purity by LC-MS and its concentration was estimated based on the amount of un-reacted PALO remaining in the sample, quantified by LC-MS using PALO as standard (Figure 24).

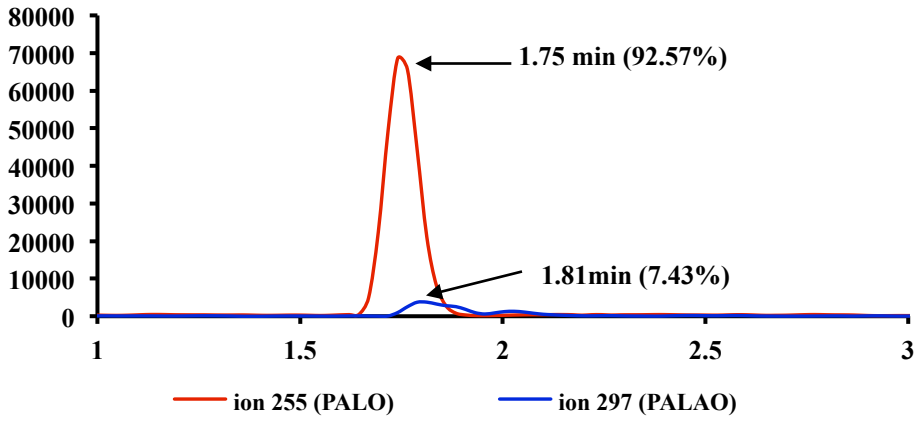


Figure 24: LC-MS analysis of chemically synthesized PALAO. The ions with an m/z corresponding to that of PALO (255m/z, Red line) and PALAO (297m/z, Blue line) were analyzed for quantification purposes. The amount of PALAO synthesized was quantified based on the amount of unreacted PALO present in the sample using known amounts of PALO as standard.

Using LC-MS analyses, the concentration of PALAO in the sample was estimated to be 9mM.

Although the purity of the chemically synthesized PALAO was not optimal we tested its ability to inhibit *Xc*AOTCase activity as presented in the next section.

Chemically synthesized PALAO inhibits XcAOTCase activity

Chemically synthesized PALAO was assayed without further purification. In order to discriminate between the effects of PALAO and the rest of the components of the reaction, each compound was tested at the higher concentration used in the synthesis.

For comparison purposes, PALO was assayed at similar concentrations as estimated for PALAO (Figure 24)

None of the components of the chemical reaction affected the enzymatic activity of *Xc*AOTCase under the conditions of the assay (data not shown) except for PALAO (Figure 24), thus proving that as hypothesized, PALO, due to the lack of acetylation at the aN-group has no effect on *Xc*AOTCase activity, while PALAO behaves as a potent inhibitor.

Based on these results the estimated half-maximal inhibition was calculated to occur at 22.4nM of PALAO.

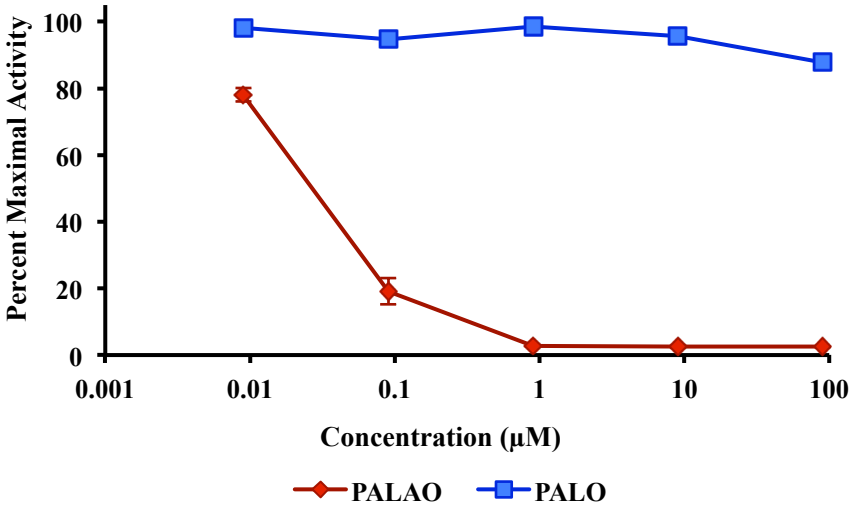


Figure 24: Effects of unpurified PALAO on *Xc*AOTCase activity. The effects of PALAO on *Xc*AOTCase activity were tested using 2µg of purified enzyme assayed in 50mM Tris-HCl pH 8.3 buffer containing 0.1mM CP and 5mM AOrn in the presence of different concentrations of PALAO and PALO at 25°C. The reactions were stopped after 5 minutes incubation with fresh color reagent (Materials and Methods). Average of three measurements and standard deviations for each point are shown.

To further characterize the inhibitory effects of PALAO, this compound was chemically synthesized and purified by IMI TAMI Institute for Research and Development, Ltd. (Haifa, Israel).

Inhibition studies using purified PALAO

The effects of purified PALAO, at 0, 0.0075, 0.01 and 0.0125mM on *Xc*AOTCase activity were tested in the presence of constant amount of CP (1mM) while the concentration of AOrn was varied from 0 to 5mM (Materials and Methods).

Although the V_{max} was only slightly affected by the presence of PALAO (Figure 25, panel A), the biggest effect was detected at the K_m level (Figure 25, panel B).

The inhibition constant (K_i) calculated for PALAO was 18.23 µM (about 23-fold less efficient than PALO for *E. coli* OTCase[75]).

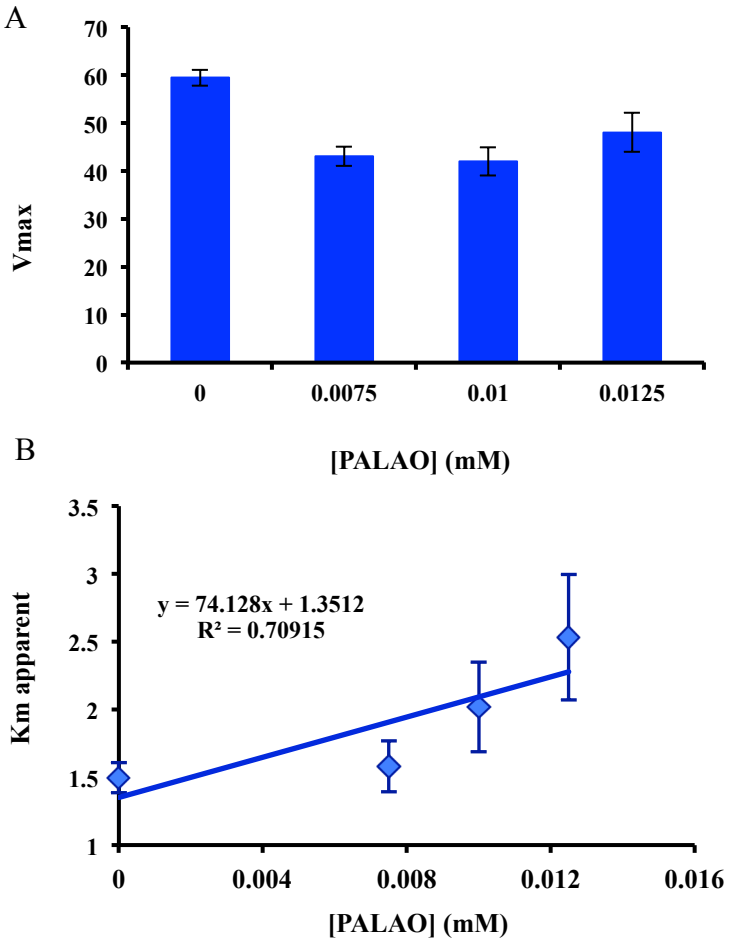


Figure 25: Effects of purified PALAO on the V_{max} and K_m of the AOTCase activity of *Xc*ArgF'. The inhibitory effects of purified PALAO were studied using 2µg of purified *Xc*AOTCase as described in the text. Panel A shows the effects on the V_{max} of the reaction relative to absence of PALAO. Panel B shows the effects on the K_m in the absence and presence of PALAO. From these results we determined that the K_i of PALAO for *Xc*AOTCase was 18.23µM.

The K_i of PALAO for *Xc*AOTCase obtained using purified PALAO, is in concordance with the estimated K_i calculated using unpurified PALAO chemically synthesized in our laboratory, 22.4µM (Figure 24), probing the high purity of the compound synthesized as well as the accurate estimation of the concentration.

2.4.4 Can polyamines affect *Xc*AOTCase activity?

There are other sources of ornithine and arginine in the cells but the implication of the arginine biosynthetic pathway as a provider of precursors for polyamines biosynthesis is evident (Figure 1). Since the arginine biosynthetic pathway is feedback controlled by the end product, it would be expected that polyamines might exert some effect over the enzymes that controls the key steps in

arginine biosynthesis, in fact *P. aeruginosa* NAGS is inhibited synergistically by polyamines[76], another examples is the catabolic OTCase from *P. aeruginosa* which is inhibited by spermidine[40, 77].

To test if *Xc*AOTCase activity could be controlled by polyamines we measured it in the presence of the polyamines putrescine and spermidine. We used AOrn saturation curves for the study of the effects of these polyamines on *Xc*AOTCase activity (Table 2).

Table 2: Effects of polyamines on *Xc*AOTCase activity. 2µg of purified *Xc*AOTCase were assayed in 50mM Tris-HCl pH 8.3 as described in Materials and Methods in the presence of several concentrations of the polyamine in study maintaining constant the CP concentration at 0.1mM while varying the AOrn concentration. Data were fitted to the Michaelis-Menten equation for enzyme kinetics for each concentration of polyamine assayed. $V_{\max}^{inh}$ and K_m^{inh} represent the V_{max} and the K_m respectively in the presence of the polyamine at each concentration

Polyamine	Concentration (mM)	$\frac{V_{\max}^{inh}}{V_{\max}^{0}}$	$\frac{K_m^{inh}}{K_m^{0}}$
Putrescine			
	0	1.00	1.00
	0.1	1.01	1.04
	1	0.95	1.01
	5	0.94	0.97
Spermidine			
	0	1.00	1.00
	0.1	0.85	0.72
	1	0.91	0.72
	5	0.93	0.60

As Table 2 shows, putrescine, under the conditions of the assay did not affect the kinetic parameters of *Xc*AOTCase and therefore its activity.

Contrary to what we were expecting, spermidine did not behave as an inhibitor, but as an activator by increasing *Xc*AOTCase affinity for AOrn.

Probably, by reducing the K_m for AOrn the polyamines biosynthetic pathway can be somehow controlled since AOrn is transcarbamylated more efficiently than deacetylated therefore regulating the availability of the precursor ornithine. Although these effects are not that significant when compared to those of arginine on NAGS activity[19]. Therefore there should be another mechanism by which the availability of precursors of polyamines is controlled in those organisms that posses an AOTCase instead of OTCase.

2.5 N-acetyl-L-citrulline continues the arginine biosynthetic pathway by means of *Xc*ArgE.

The fact that *Xc*ArgF' is indeed an AOTCase and therefore the product of the reaction catalyzed by this enzyme is N-acetyl-L-citrulline as demonstrated before,

raises a question of biological importance: Can N-acetyl-L-citrulline continue the arginine biosynthetic pathway?

Initial complementation studies performed by Dr. Malamy suggested that N-acetyl-L-citrulline could be deacetylated to form citrulline thus continuing the arginine biosynthetic pathway as described by the following equation:

$$\text{N-acetylcitrulline} + H_2O \xrightarrow{x_c\text{ArgE}} \text{L-citrulline} + \text{acetate}$$

This reaction should be catalyzed by x_cArgE. To test this hypothesis we used purified x_cArgE (AOD, Figure 2) generously provided by Dr. Daniel Shi and we assayed its activity in the presence of AOrn and ACit.

In this case, 20.6µg of purified x_cArgE were incubated for 10 minutes at 37°C in 80mM potassium phosphate buffer pH 7.0 containing 0.02mM $CoCl_2$, 0.08mM glutathione (reduced) and either AOrn or ACit in a final volume of 500µl. After 10 minutes the reaction was stopped by addition of 500µl of 7% sulfosalicylic acid. After protein precipitation by centrifugation samples were loaded in an amino acid analyzer (Beckman) for its quantification using citrulline and ornithine as standards. The results obtained are represented in Figure 26.

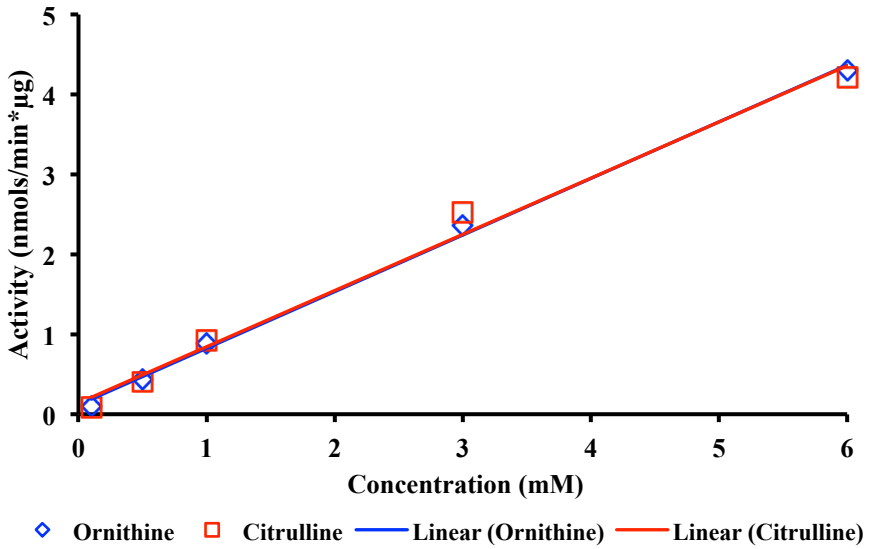


Figure 26: Deacetylation of AOrn and ACit catalyzed by x_cArgE. The activity of x_cArgE to deacetylate either AOrn or ACit was assayed using 20.6µg of purified enzyme in 80mM potassium phosphate buffer pH 7.0 in the presence of 0.02mM $CoCl_2$, 0.08mM glutathione (reduced) and either AOrn or ACit in a final volume of 500µl. After an incubation of 10 minutes at 37°C the reaction was stopped with 500µl of 7% sulfosalicylic acid and the samples were loaded in a Beckman amino acid analyzer after protein precipitation (by centrifugation). Activity is expressed as nmols of citrulline (red squares) or ornithine (blue diamonds) per minute and per µg of protein for each concentration assayed.

These results support the hypothesis that ACit continues the arginine biosynthetic pathway in *Xanthomonads* through deacetylation catalyzed by ArgE.

Furthermore, these results provided key information to elucidate the arginine biosynthetic pathway in these organisms.

The fact that *Xc*ArgE is able to catalyze the deacetylation of AOrn as well as ACit can be used by these organisms (those provided with an AOTCase instead of an OTCase) to regulate the levels of ornithine thus controlling the arginine and the polyamines biosynthetic pathways.

The characterization of *Xc*ArgE and the possible regulatory mechanisms that control its activity is an interesting continuation of this work, unfortunately is beyond the scope of the present Thesis work and therefore has not been studied.

2.6 Summary

The results presented in this chapter: **Functional characterization of a novel type of transcarbamylase that uses acetylornithine, identified in *Xanthomonas campestris pv campestris***, demonstrates for the first time that there is a new member of the transcarbamylases family of enzymes which catalyzes the carbamylation of N-acetylornithine to produce N-acetylcitrulline and therefore denominated acetylornithine transcarbamylase (AOTCase).

In the first part of this chapter I have provided conclusive evidences proving the substrates of the reaction catalyzed by *Xc*ArgF' to be carbamylphosphate and aN-acetyl-L-ornithine, and that the product of the reaction is aN-acetyl-L-citrulline.

The specificity of the reaction has been demonstrated and AOrn is the only substrate that *Xc*AOTCase can utilize. Additionally the optimal conditions for the assay of the activity of *Xc*AOTCase have been determined.

The kinetics parameters of the activity catalyzed by *Xc*AOTCase have been determined using saturation curves for both substrates (CP and AOrn) and are summarized in Table 3.

Table 3: Kinetic parameters determined for *Xc*AOTCase. Summary of the kinetic parameters determined for the AOTCase from *X. campestris pv campestris*.

Substrate	Km (mM)	Vmax (µmols/min*µg)
CP	0.01 ± 0.0007	50.68 ± 0.48
AOrn	1.05 ± 0.23	65.05 ± 4.77

Although the kinetic parameters shown in Table 3 that characterizes the activity of *Xc*AOTCase are on the range of those calculated for other OTCases[41, 68-71], the affinity of *Xc*AOTCase for CP is unusually high, being comparable only to that of *S. cerevisiae* OTC[68] (0.019mM).

On the contrary to other members of the transcarbamylases family like OTC and ATC that are inhibited by their substrates, ornithine and aspartate respectively[41, 72-74], *Xc*AOTCase is not inhibited by the substrate AOrn, even at concentrations as high as 50mM (Figure 21). This fact lead us to think that *Xc*AOTCase is not a crucial point of control for the metabolic flow of intermediates in the arginine biosynthetic pathway, although the lack of this activity renders arginine auxotrophy, highlighting the importance of the reaction catalyzed by

*Xc*AOTCase.

As indicated before (Section 2.5), regulation of *Xc*ArgE, both at the transcriptional and enzymatic levels could be the mechanism by which, the organisms that possess an AOTCase, controls the metabolic flow of intermediates between arginine and polyamines biosynthesis.

The results of the studies with polyamines showed very little or no effect on *Xc*AOTCase activity, corroborating the hypothesis that *Xc*ArgE is a most likely point of control for the arginine and polyamines biosynthetic pathways.

In this chapter I have also proven that acetylation of the aN-group is not only necessary for substrate recognition but that is also necessary for certain compounds to exert an inhibitory effect on *Xc*AOTCase activity. This has been demonstrated not only by the lack of enzymatic activity when ornithine was used as substrate but by the lack of inhibition when *Xc*AOTCase was assayed in the presence of norvaline or PALO while the acetylated versions of each of these compounds were able to inhibit the enzymatic activity of *Xc*AOTCase.

Chapter 3. Demonstration that a structurally obvious transcarbamylase from *Bacteroides fragilis*, of previously unknown specificity catalyzes the transcarbamylation of succinylornithine. Functional characterization of this transcarbamylase.

3.1 *Bacteroides fragilis*

Bacteroidetes is one of the major lineages of bacteria that arose early during the evolutionary process. *Bacteroides* species are anaerobic, bile-resistant, non-spore-forming, Gram-negative rods. Its taxonomy has undergone major revisions in the last decades and contains >20 species, many of these were isolated as single strains from human feces[78]. The numbers of *B. fragilis* isolates are 10 to 100-fold lower than those of other intestinal *Bacteroides* species, yet *B. fragilis* is the most frequent isolate from clinical specimens and is regarded as the most virulent *Bacteroides* species.

Recent genomic and proteomic advances have greatly facilitated our understanding of the uniquely adaptive nature of *Bacteroides* species. Thus far (at the time of writing the present work) the genomes of *B. fragilis*, *B. thetaiotaomicron*, *B. vulgatus* and *Parabacteroides distasonis* have been sequenced and an increasing number is in the process of being completed. The first to be sequenced were *B. fragilis* and *B. thetaiotaomicron*. Both genomes share a few interesting facts; there is an unusually low gene content for their genome size, which reflects a large number of proteins containing >1000 amino acids, many of these predicted proteins were assigned putative functions based on homology with known bacterial proteins (~60% in *B. thetaiotaomicron*). A second common feature in both species is the presence of extensive DNA inversions that may control the expression of a large number of genes. Also, both species exhibit multiple paralogous groups of genes, i.e., genes that seem to have derived from a common ancestral gene and have since diverged from the parent copy by mutation and selection of drift[78].

Bacteroides species has established throughout the evolution a relationship that can be considered mutualism since both the bacteria and the human host experience increased fitness as a result of the relationship[79]. Is when the *Bacteroides* organisms escape the gut, usually resulting from rupture of the gastrointestinal tract of intestinal surgery, they can cause significant pathology[78] with high morbidity and mortality rates.

Although *B. fragilis* accounts for only 0.5% of the human colonic flora[80], it is the most commonly isolated anaerobic pathogen, due in part to its potent virulence factors[78], which make this a very clinically important microorganism.

Arginine is an essential amino acid for the survival of *B. fragilis* as shown by Dr. Malamy using transposon-mediated mutagenesis. As indicated in the Introduction the lack of *argF'* in *B. fragilis* rendered arginine auxotrophy highlighting the importance of this gene in the arginine biosynthetic pathway for these organisms.

Even though the crystal structure of the protein encoded by *Bf*a*rgF'* was solved[47], its activity remained unknown.

In this chapter I will present evidences proving that *Bf*ArgF' catalyze the transcarbamylation of N-Succinyl-L-ornithine and therefore is a new member of the transcarbamylases family.

3.2 Identification of potential substrates of *Bf*ArgF'

The identification of the AOTCase activity of *Xc*ArgF' implied the possibility that, at least, some of the transcarbamylases of unknown activity identified in the studies that preceded this thesis could use different derivatives of ornithine as substrates for the reaction, since, in contrast to the CP-binding site, that is structurally conserved across *Xc*AOTCase, *Bf*ArgF', in human OTCase and *E. coli* ATCase[63], the putative binding site for the second substrate is strikingly different from that of other known transcarbamylases[47] (Figure 27).

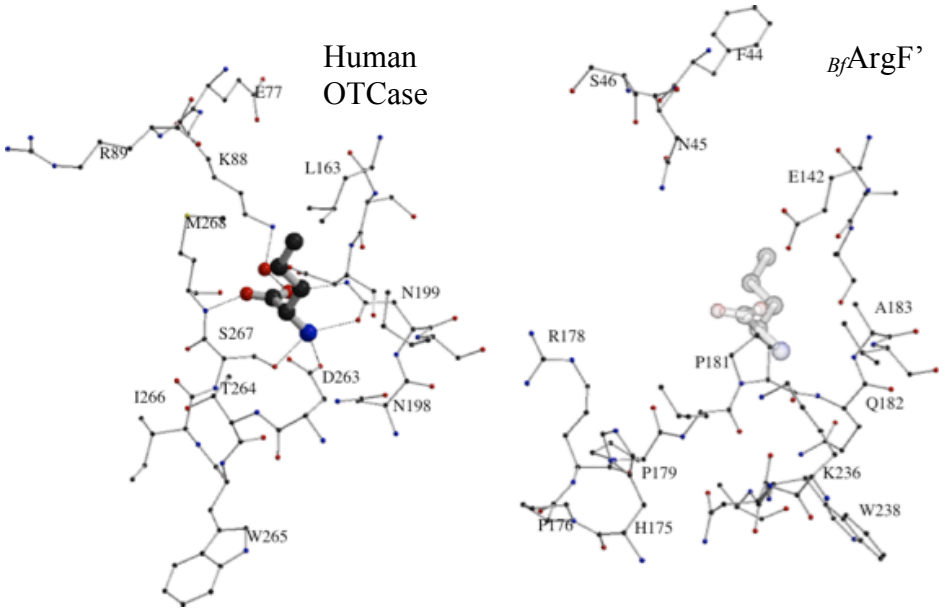


Figure 27: Comparison of the ornithine binding domain in human OTCase and the putative ornithine-binding domain in *Bf*ArgF'. Schematic representation of the ornithine binding domains in human OTCase (left) and *Bf*ArgF' (right). Residues involved in substrate recognition are indicated by single letter code and residue ID number. Putative localization of ornithine in the *Bf*ArgF' protein is represented in shadows (left).

As shown in Figure 15 *Bf*ArgF' lacks the ability to transcabamylate AOrn to produce ACit, therefore presenting a new challenge; If this enzyme is required for the *de novo* arginine biosynthesis in *B. fragilis* but can not use neither ornithine nor acetyl ornithine, what it is the substrate of the reaction that can continue the pathway?

Could it be N-Succinyl-L-ornithine?

N^2-succinyl-L-ornithine (SOrn) is one of the intermediates in the catabolic arginine succinyltransferase pathway but the enzyme that utilizes this intermediary (astC) has very little similarities with *Bf*ArgF' (13% identity and 21.7% similarity). Based on these similarities there is no evidence that SOrn could be the substrate for *Bf*ArgF'.

Although when the structures of L-ornithine, AOrn and SOrn are compared (Figure 28) it seems to be a good candidate since its structure is more voluminous than that of AOrn and could potentially fit the binding pocket for the second substrate in *Bf*ArgF'.

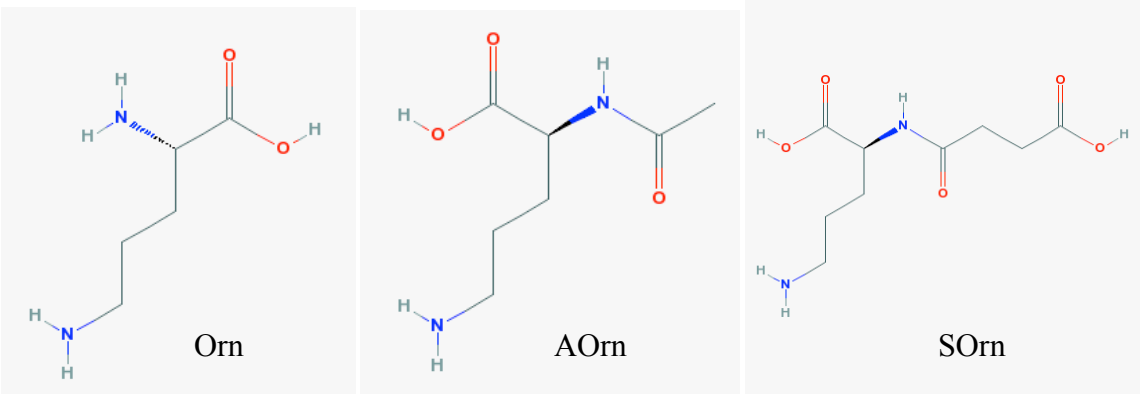


Figure 28: Schematic representation of the substrates for OTCase, AOTCase and the putative substrate for *Bf*ArgF'. Ornithine, the substrate of OTCase is shown on the left panel of the figure, acetylornithine, the substrate of AOTCase is shown in the middle panel and succinylornithine is shown on the right panel of the figure. These schematic representations illustrate the structural differences between these compounds.

3.3 Chemical synthesis of the potential substrate N-Succinly-L-ornithine and initial activity assays

To test this hypothesis I synthesized SOrn chemically as follows:

Initially, 1.076mmols of *t*-Boc-Ornithine (ornithine with the α-amino group temporarily protected by a tert-butyloxycarbonyl group) were dissolved in 5ml of water. The pH of the reaction was kept neutral with 40% (w/v) NaOH while 1.076mmols of succinic anhydride were added. Once the addition of succinic anhydride was completed and the pH was neutral the reaction was considered finalized.

Without further analyses the product of the reaction (presumably SOrn in its majority) was tested as a substrate of *Bf*ArgF'.

Initial activity assays to test SOrn

To determine whether or not *Bf*ArgF' could utilize SOrn as substrate we used the chemically synthesized SOrn.

In this experiment, *E. coli* C6.1 cells containing the plasmid pTrc99a-BfArgF-His were induced with IPTG to over express *Bf*ArgF' as indicated in Materials and Methods and then lysed in Ni-column purification buffer A (Materials and Methods), were assayed using the colorimetric method for OTC and AOTCase activities.

The cells lysate was assayed without further purification in 50mM Tris-HCl pH 8.3 buffer containing 4.8mM CP and an unknown amount of SOrn (a ½ dilution

of the chemically synthesized SOrn without further purification) (Figure 29). Boiled cells lysate and AOrn were used as negative controls.

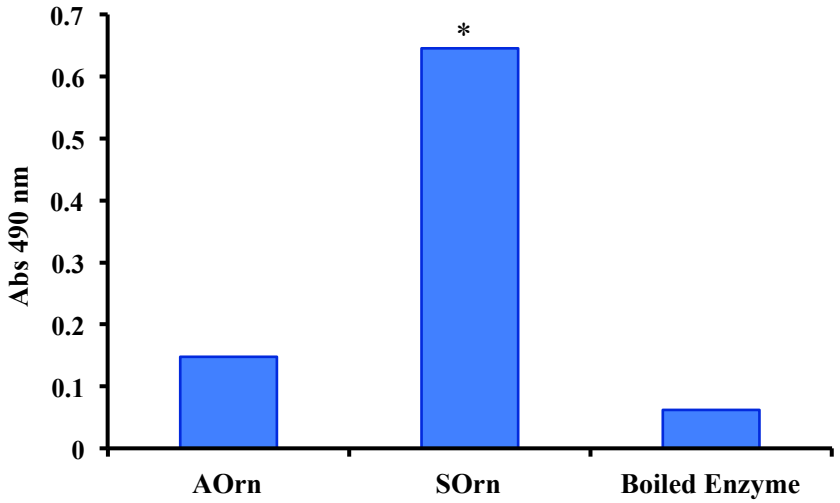


Figure 29: First attempt to elucidate the nature of the second substrate of the reaction catalyzed by *Bf*ArgF'. 20µl of cells lysate were assayed in 50mM Tris-HCl pH 8.3 buffer in the presence of 4.8mM CP and an unknown amount of SOrn or 5mM AOrn, for 5 minutes at 25°C. (*) Sample absorbance was above the detection limit of the instrument (spectrophotometer).

Although the amount of color obtained for the sample assayed with SOrn was over the detection limit of the spectrophotometer used for the quantification (≥3 units of absorbance) and, therefore, the result was not accurate, this experiment showed for the first time, that the activity of *Bf*ArgF' seems to be that of N-succinyl-L-ornithine transcarbamylase (SOTCase).

In this experiment the cell lysate contained some activity with AOrn that could be explained by the presence of other enzymes in the sample assayed.

While we could not rule out the possibility of an artifact in the SOTCase activity obtained due to the induction of undesired pathways and proteins when inducing the expression of *Bf*ArgF', the boiled enzyme control showed that the sample did not contain a component that could be responsible for the absorbance and thus the compound formed when using SOrn as substrate is most likely the product of the activity of the over expressed *Bf*ArgF'.

To confirm that *Bf*ArgF' indeed, catalyzes the conversion of SOrn to SCit in the presence of CP, *Bf*ArgF' was purified to homogeneity as described in Materials and Methods and assayed in the presence of CP and chemically synthesized SOrn.

The SOTCase activity was tracked through the purification process to identify the fraction that contained such activity (Figure 30). For this experiment 2 µg of protein from each fraction were assayed in the presence of 1 mM CP and a ½ dilution of the chemically synthesized SOrn for 10 minutes using the 1 hr OTCase assay as described in Materials and Methods.

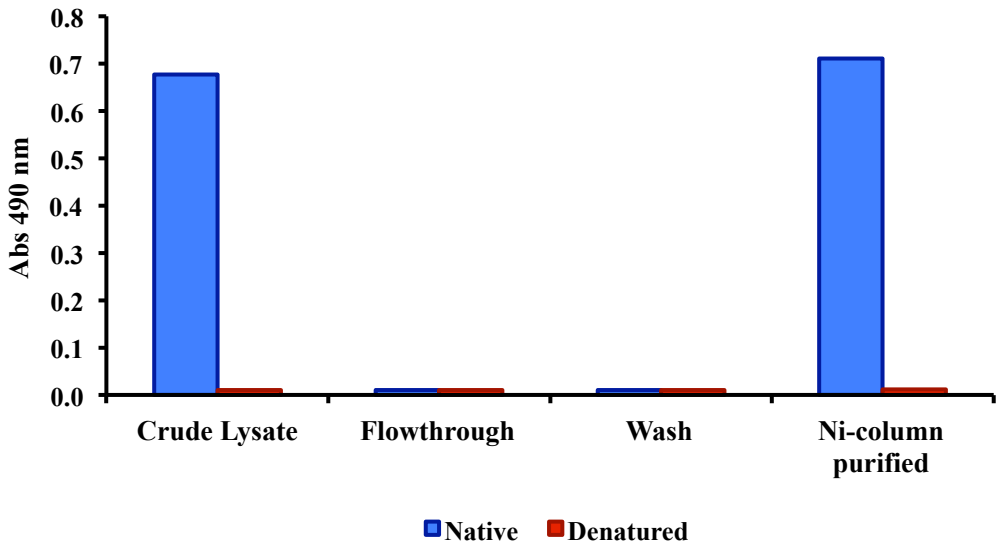


Figure 30: SOTCase activity levels in the fractions collected during the purification of *Bf*ArgF'. 2µg of protein from each of the fractions collected during the purification of *Bf*ArgF' were assayed for SOTCase activity. The fractions were: Crude Lysate: cells lysated in Ni-column purification buffer A prior to be loaded in the column; Flowthrough: fraction collected after the cells lysate has been loaded onto the column, containing a mixture of proteins that do not bind the Ni-column; Wash: fraction containing proteins eluted using mild stringent conditions (80% buffer A, 25% buffer B, Materials and Methods); Ni-column purified: containing purified BfArgF' after elution from the column using stringent conditions (50% buffer B). Measurements were performed in duplicate for each fraction and the average is represented in the graph.

This experiment shows that the activity is restricted to the fractions that contains the over expressed protein, that is the crude lysate and the Ni-column purified fractions, proving that *Bf*ArgF' catalyzes the transcarbamylation of SOrn to produce SCit in the presence of CP.

3.4 Identification of N-succinyl-L-Citrulline as the product of the reaction catalyzed by *Bf*ArgF'

Although the experiments described in the previous sections demonstrate that the substrate of the reaction catalyzed by *Bf*ArgF' is SOrn, they only indicate that the product of such reaction contains an ureido group as described in section 2.2.1, Chapter 2. It is expected that being CP and SOrn the substrates of the reaction, the product should be SCit, which contains an ureido group and through de-succinylation can continue the arginine biosynthetic pathway in these organisms.

In order to prove this hypothesis both, chemically and enzymatically synthesized SCit were analyzed by NMR and LC-MS (Figure 31) as specified in Materials and Methods.

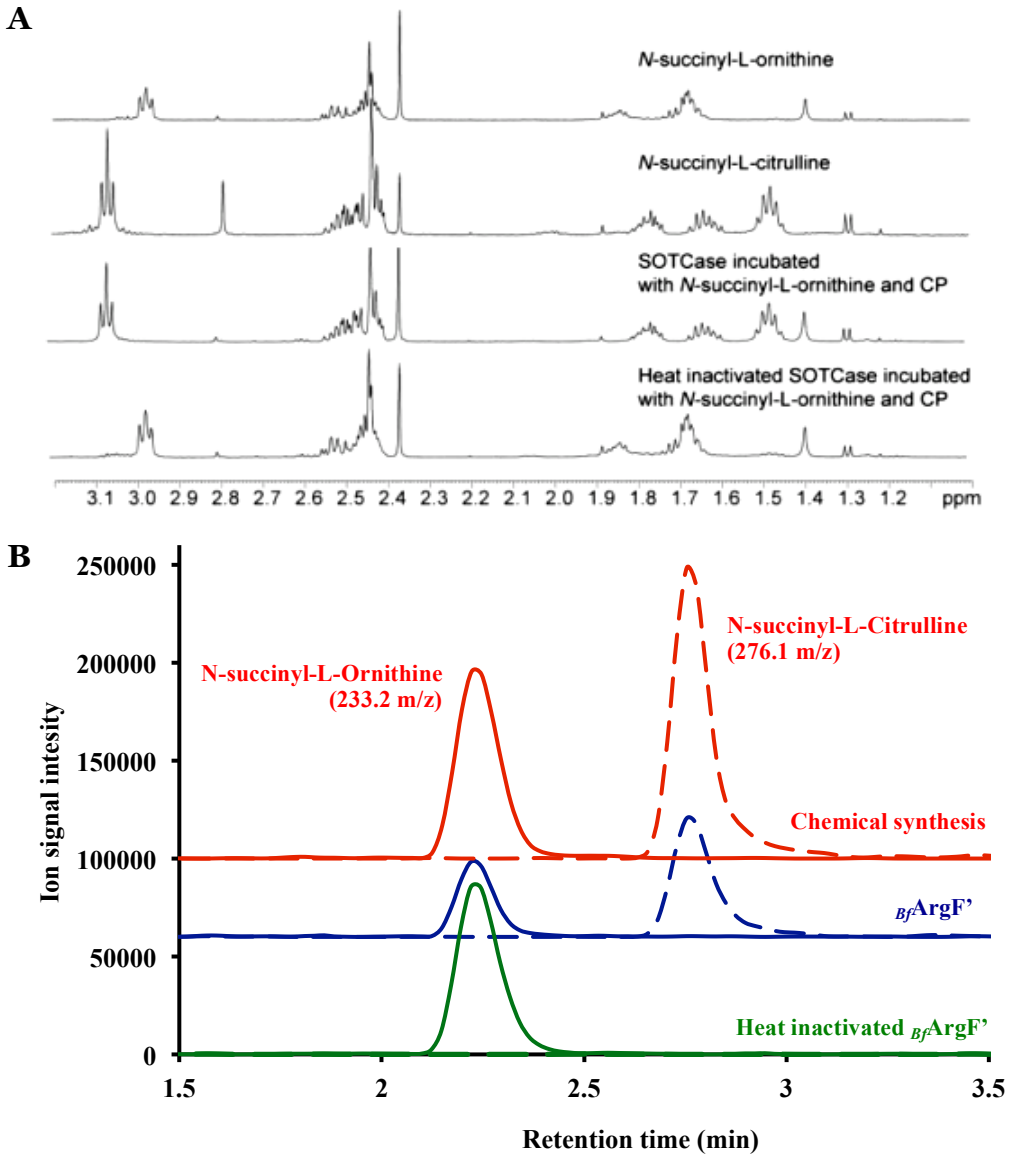


Figure 31: NMR and LC-MS analysis of the product of the reaction catalyzed by *Bf*ArgF'. **A**, proton NMR of chemically synthesized N-succinyl-L-ornithine and N-succinyl-L-citrulline are shown compared with the products formed using CP and SOrn incubated with either *Bf*ArgF' or heat inactivated *Bf*ArgF'. **B**, The enzymatic product of the SOTCase reaction (dashed blue line) analyzed using LC-MS has the same retention time and m/z as the chemically synthesized SCit (dashed red line). Ions corresponding to SOrn and SCit were extracted from the total spectra of chemically synthesized compounds and the enzymatic reaction, and then plotted. Heat inactivated *Bf*ArgF' used as negative control. Base lines have been offset vertically for clarity.

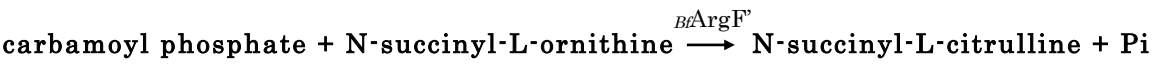
The peaks seen in the proton NMR spectra of the enzymatically-generated product have chemical shifts that are identical to those of the chemically synthesized SCit (Figure 31 panel A).

Thick lines in Figure 31, panel B, represent SOrn and dashed lines SCit. In red are shown the extracted m/z ions for a concentration of 5 mM of both compounds

chemically synthesized run independently. In blue lines the extracted m/z ions for both compounds after incubation of *Bf*ArgF' in the presence of 5 mM SOrn and 1 mM CP, both peaks are present but the one corresponding to SOrn has decreased with respect to the initial concentration (same as in the chemically synthesized run), indicating that this substrate has been consumed during the reaction. And in green lines the same ions extracted from a sample in which *Bf*ArgF' has been heat inactivated and incubated as the native sample, where only the peak for SOrn is present and therefore, no SCit is formed, as expected. As this Figure 31 shows, the retention times and the m/z of the enzymatically-synthesized product are identical to those of the chemically synthesized SCit.

The results obtained by these two techniques confirmed that the product of the reaction catalyzed by *Bf*ArgF' is N-succinyl-L-citrulline.

Based on the results presented in the previous sections we concluded that the activity of *Bf*ArgF' is that of N-succinyl-L-ornithine transcarbamylase (SOTCase) and would be described by the following equation:



This SOTCase activity has not been described previously and therefore *Bf*SOTCase can be added as a new member of the transcarbamylases family.

3.5 Determination of the extinction coefficients for citrulline, acetylcitrulline and succinylcitrulline

Due to the lack of chemically synthesized and purified SCit, thus far we used ACit as standard for the quantification of the product formed by *Bf*SOTCase.

We observed that the same concentration of Cit and ACit produce different amount of color under the same conditions, therefore we expected that replacing the acetyl group for the succinyl group would have similar effects on the extinction coefficient.

In order to accurately characterize the SOTCase activity of *Bf*ArgF', SOrn and SCit were synthesized and purified by Chiral Quest Co. (Monmouth Junction, NJ).

We used purified ACit and SCit to determine the extinction coefficients for each compound.

We used a colorimetric assay that detects the ureido group present in all three compounds, to determine the extinction coefficients. Samples containing 0, 5, 10, 25, 50, 75 and 100 nmols of either Cit, ACit or SCit in 1ml of 50mM Tris-HCl pH 8.3 were treated with 1ml of fresh color reagent (Materials and Methods) and incubated at room temperature over night under constant shaking (100rpm) and protected from the light. Following the over night incubation at room temperature, samples were incubated at 42°C in a water bath under light for color development completion. After that samples were placed on ice water and the absorbance was measured at 466nm using a spectrophotometer.

We applied the Beer-Lambert law that is characterized by the equation, to determine the extinction coefficients:

$$A = \varepsilon lc$$

where A represents the absorbance at any given wavelength, e represents the molar absorptivity or extinction coefficient; *l* is the cell path length and *c* the analyte concentration.

The results obtained in this experiment are shown in Figure 32.

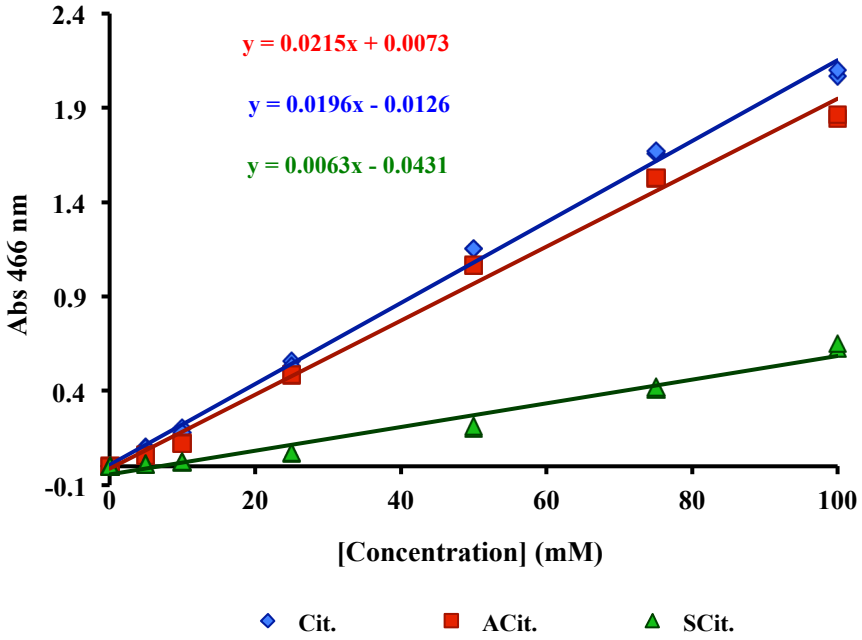


Figure 32: Determination of the extinction coefficients for citrulline, acetylcitrulline and succinylcitrulline. Each concentration was measured in independent duplicates. Both measurements for each concentration are shown in the graph to highlight the accuracy of the determination. Extinction coefficients can easily be calculated from the slopes of these representations when using linear regression fitting of the data

The extinction coefficients calculated from these results were: 0.0215, 0.0196 and 0.0063 L mol^{-1} cm^{-1} for Cit, ACit and SCit respectively. These extinction coefficients are quite different highlighting the importance of the use of the appropriate reference compound when determine enzyme activity and calculating kinetic parameters.

Purified SCit and SOrn were used as reference and substrate respectively, to accurately determine the kinetic parameters of *Bf*SOTCase.

3.6 Initial characterization of the SOTCase activity of *Bf*ArgF'

SOrn saturation curve

For the determination of the K_m for SOrn and the V_{max} of *Bf*SOTCase, 0.2µg of purified enzyme were assayed in 50mM Tris-HCl buffer pH 8.3 in the presence of a constant amount of 4.8 mM CP while the SOrn concentration was varied from 0 to 25mM, in a final volume of 1ml (Materials and Methods). After 5 minutes incubation at 25°C the reaction was stopped by the addition of freshly made color reagent (Materials and Methods). Each point of the saturation curve was measured in triplicate (Figure 33).

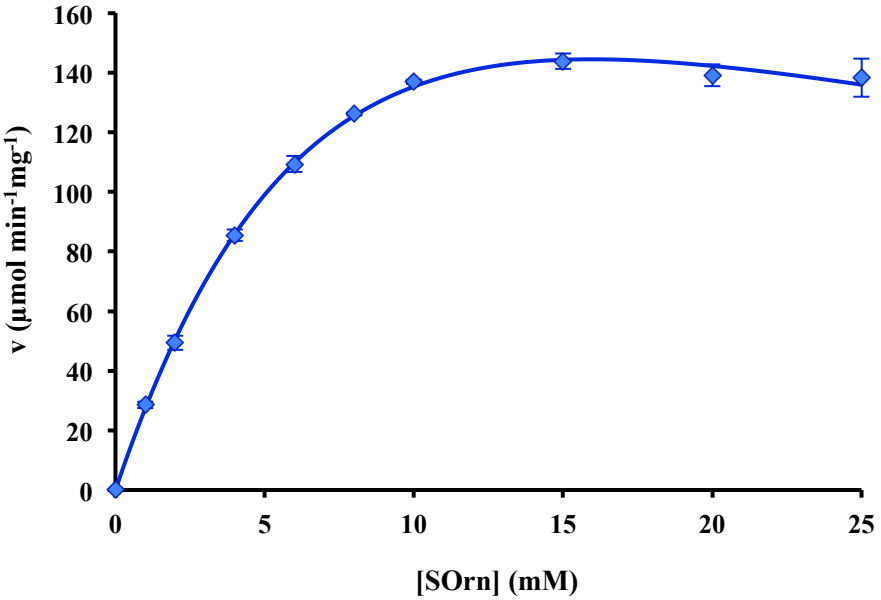


Figure 33: Determination of the kinetics parameters K_m and V_{max} for SOrn of *Bf*SOTCase. 0.2µg of purified *Bf*SOTCase were assayed in 50mM Tris-HCl pH8.3 buffer in the presence of 4.8mM CP while varying the SOrn concentration between 0 and 25mM in a final volume of 1ml. After 5 minutes incubation at 25°C the reaction was stopped and the product formation was quantified using SCit as standard. Data showed slight substrate inhibition above 15mM of SOrn. The K_m and V_{max} values determined using the substrate inhibition equation to fit the data were: 12.5±1.5mM and 370.6±32.6µmols $min^{-1}mg^{-1}$ respectively, while the calculated K_{si} was 20.4±3.3mM.

Under these conditions the enzyme shows a slight substrate inhibition at concentrations of SOrn above 15mM. Data were fitted to the substrate inhibition equation (described in section 3.7.1).

The kinetic parameters calculated for SOrn in this experiment were:

V_{max}= 370.6 ± 32.6 µmol $min^{-1}mg^{-1}$

K_m= 12.5 ± 1.5 mM

K_{si}= 20.4 ± 3.3 mM

The K_m of *Bf*SOTCase for the substrate SOrn is the highest compared to other members of the transcarbamylase family, especially when compared with human and *S. cerevisiae* OTCases[41, 68]).

CP saturation curve

The kinetics parameters K_m and V_{max} of *Bf*SOTCase for the substrate CP were determined using 0.2µg of purified enzyme as previously described. *Bf*SOTCase was assayed in 50mM Tris-HCl buffer pH 8.3 in the presence of a constant amount of 20mM SOrn (approximately 2 times the previously calculated K_m, since higher concentrations would inhibit the enzyme) while varying the concentration of CP from 0 to 2mM. Figure 34 shows the CP saturation curve obtained under these conditions.

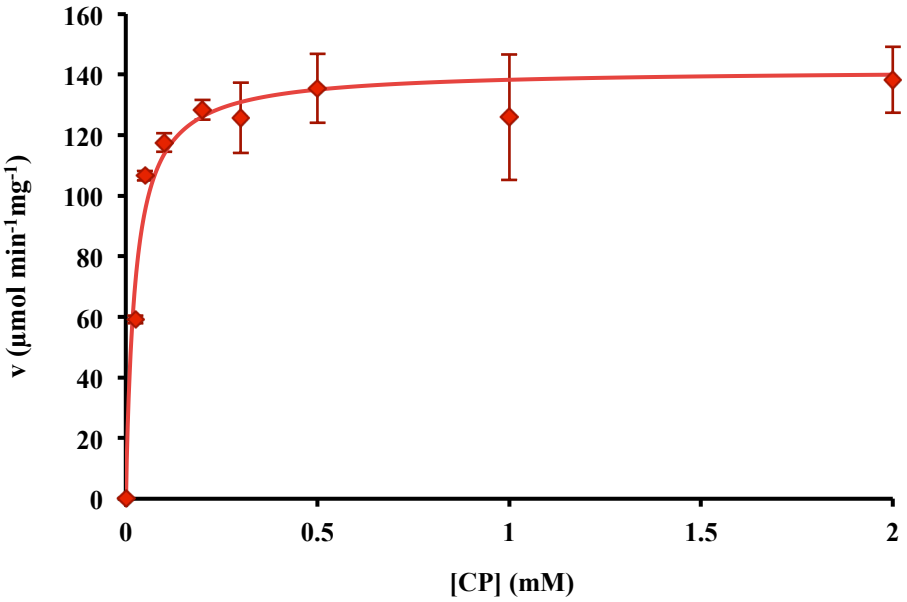


Figure 34: Determination of the kinetic parameters K_m and V_{max} for CP of *Bf*SOTCase. 0.2µg of purified *Bf*SOTCase were assayed in 50mM Tris-HCl buffer pH 8.3 containing a constant amount of SOrn of 20mM while varying the CP concentration from 0 to 2mM. After 5 minutes incubation at 25°C the reaction was stopped and the product formation was quantified using SCit as standard. Data were fitted to the Michaelis-Menten equation for enzyme activity using the program GNUPLOT (by Williams & Kelley, 2004), and the K_m and V_{max} were determined to be 0.025 ± 0.003mM and 141.7 ± 3 µmol $min^{-1}mg^{-1}$ respectively.

Data were fitted to the Michaelis-Menten equation for enzyme kinetics using the software GNUPLOT and the kinetic parameters K_m and V_{max} were determined to

be 0.025 ± 0.003mM and 141.7 ± 3 μmol $min^{-1}mg^{-1}$ respectively.

The affinity of *Bf*SOTCase for CP is quite high as in the case of the affinity of *Xc*AOTCase[61] for CP and is only comparable to that of *S. cerevisiae* OTCase[68].

These results show that *Bf*SOTCase behaves as other members of the transcarbamylases family previously described and s in the case of *Xc*AOTCase, *Bf*SOTCase is a highly specific enzyme unable to catalyze the transcarbamylation of any substrate other than SOrn as demonstrated before (Figures 15 and 29).

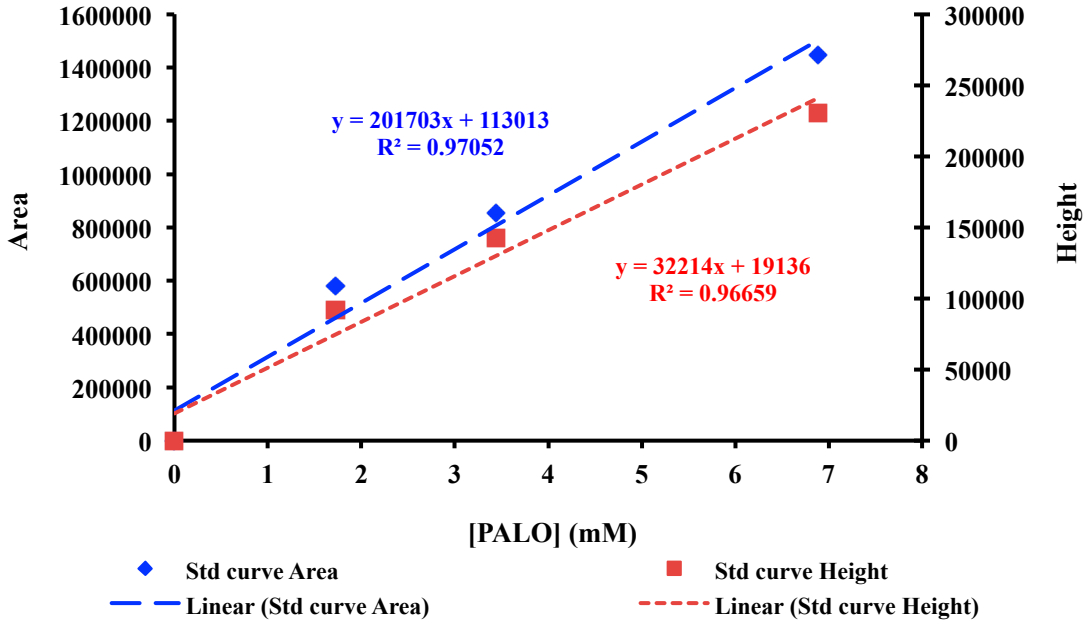
As demonstrated for *Xc*AOTCase, the acetyl group was required both for activity as well as for inhibition (Figures 15, 22, and 24), therefore it was expected that the succinylated version of compounds such as PALO would exert inhibitory effects over *Bf*SOTCase since PALO and PALAO are two of the most potent inhibitors of OTCase and AOTCase respectively[61, 75].

To test this hypothesis I chemically synthesized the succinylated derivative of PALO: αN-succinyl-δN-phosphonoacetyl-L-ornithine (PALSO).

3.7 Chemical synthesis of the potential *Bf*SOTCase inhibitor αN-succinyl-δN-phosphonoacetyl-L-ornithine (PALSO).

For the chemical synthesis of PALSO 5.1mg (0.02mmols) of PALO were dissolved in 200μl of water. The pH of the solution was kept neutral with 40% (w/v) NaOH while 0.2mmols of succinic anhydride were added at room temperature with constant stirring. Once the pH stabilized the reaction was considered completed.

The amount of PALSO chemically synthesized was estimated by LC-MS, using PALO as standard, based on the amount of un-reacted PALO present in the sample once the reaction was stopped. The amount of un-reacted PALO was estimated based on both the area and the height of the ion 296 m/z when extracted from the LC-MS results (Figure 35).



Parameter	Measurement	Remaining PALO concentration (mM)	Estimated PALSO concentration (mM)	Estimated % of conversion
Area	184406	0.177	6.708	97.4%
Height	33411.3	0.222	6.663	96.8%

Figure 35: Estimation of the concentration of chemically synthesized PALSO. The amount of PALSO was calculated by LC-MS based on the remaining un-reacted PALO present in the sample after the chemical reaction was stopped. Quantification was determined by measuring both the area and the height of the peak corresponding to the PALO ion (296.1 m/z). Bottom table shows the results obtained for the un-reacted PALO present in the sample of the chemically synthesized PALSO as well as the estimated PALSO concentration and the % of conversion from PALO to PALSO.

As shown in Figure 35, the degree of conversion was near 100%. Although both measurements (area under the peak and peak height) provided similar results for the estimation of the amount of PALSO present in the sample, we used the estimation based on the area under the peak parameter for our assays.

Chemically synthesized PALSO was used without further purification to study its effects on *Bf*SOTCase activity.

3.8 Study of the effects of PALSO on *Bf*SOTCase activity

For these assays 0.2µg of purified *Bf*SOTCase were assayed in 50mM Tris-HCl buffer pH 8.3 and 5mM SOrn and 0.1mM CP in the presence of different amounts of PALSO (ranging from 0 to 670.8µM). After 5 minutes incubation at 25°C, the reaction was stopped with fresh color reagent (Materials and Methods) and the

product was quantified using SCit as standard (Figure 36).

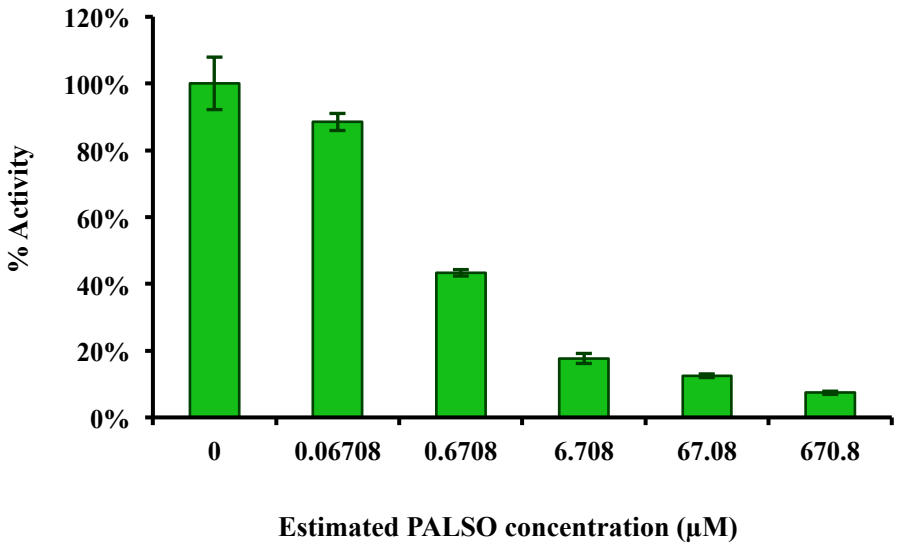


Figure 36: Inhibitory effects of PALSO on *Bf*SOTCase activity measured at saturating conditions. 0.2µg of purified *Bf*SOTCase were assayed in 50mM Tris-HCl buffer pH 8.3 containing 5mM SOrn and 0.1mM CP in the presence of different amounts of PALSO (ranging from 0 to 670.8µM). Reactions were carried out for 5 minutes at 25°C and product formation was quantified using SCit as standard.

Data were fitted to the four parameter logistic model (4PL), also called the Hill-Slope model, described by the following equation:

$$y = bot + \frac{(top - bot)}{1 + \left(\frac{x}{IC_{50}}\right)^{slope}}$$

where *y* is the percent activity and *x* is the corresponding concentration. The fitted IC_{50} parameter is the relative IC_{50}, and is defined as the concentration giving a response half way between the fitted top and bottom of the curve (http://www.ncgc.nih.gov/guidance/section3.html).

As shown in the Figure 36, PALSO is a very potent inhibitor of *Bf*SOTCase with a calculated IC_{50} of 0.45 ± 0.05 µM.

The effects of un-purified PALSO on *Bf*SOTCase activity are comparable to those of PALAO on *Xc*AOTCase activity[61] (Figure 24) and PALO on *Ec*OTCase activity[75], although the calculated IC_{50} is higher in the case of PALSO. These differences could be explained by different interactions of each bisubstrate analog to the corresponding enzyme.

3.9 Summary

In this Chapter 3: **"Demonstration that a structurally obvious transcarbamylase from *Bacteroides fragilis*, of previously unknown activity catalyzes the transcarbamylation of succinylornithine. Functional characterization of this transcarbamylase"** I have provided evidences probing that the activity of the enzyme codified by the gene *argF'* in *B. fragilis* is that of an N-succinyl-L-ornithine transcarbamylase (SOTCase).

*Bf*ArgF' specifically catalyzes the transcarbamylation of succinylornithine in the presence of CP as the specificity tests performed shown in this, as well as, in the previous chapter have demonstrated. NMR and LC-MS analyses confirmed that the product of the reaction catalyzed by *Bf*ArgF' is N-succinylcitrulline.

In this chapter the initial characterization of the reaction catalyzed by *Bf*SOTCase is presented. The kinetics parameters (K_m and V_{max}) obtained for each of the substrates of this enzyme are summarized in Table 4.

Table 4: Initial characterization parameters determined for *Bf*SOTCase. The enzyme showed slight substrate inhibition for SOrn but not with CP, under the conditions of the assay.

	SOrn	CP
K_m (mM)	12.5 ± 1.5	0.025 ± 0.003
V_{max} (μmol $min^{-1}mg^{-1}$)	370.6 ± 32.6	141.7 ± 3
K_{si} (mM)	20.4 ± 3.3	n/a

*Bf*SOTCase showed substrate inhibition by SOrn, that was not detected in *Xc*AOTCase, although it is much less pronounced than in other members of the transcarbamylases family[41, 72-74].

Based on the specificity of the reaction catalyzed by *Bf*SOTCase I have synthesized the bisubstrate analog PALSO from PALO and characterized its inhibitory effects on *Bf*SOTCase. Using PALSO chemically synthesized (97.4% pure), the IC_{50} was determined to be 0.45μM probing that PALSO is a potent inhibitor of *Bf*SOTCase activity, although its behavior is still undetermined since no purified source of PALSO is available. These results suggest that the succinylated derivatives of common inhibitors of OTCase, like norvaline, should inhibit the activity of *Bf*SOTCase, as the acetylated derivatives inhibit the *Xc*AOTCase activity.

The discovery of the SOTCase activity of *Bf*ArgF' described here, demonstrates the presence of a succinylated intermediate in the arginine biosynthetic pathway in this organism since the lack of *argF'* renders arginine auxotrophy. The fact that succinylornithine is an obligate intermediate in the arginine biosynthetic pathway implies that the initial part of this pathway utilizes succinylated precursors.

To probe this hypothesis I will elucidate the activity of the initial enzymes in the arginine biosynthetic pathway of *B. fragilis*, that is ArgA and ArgB.

The results obtained from these studies are presented in the following Chapter 4.

Chapter 4. The use of N-succinyl-L-glutamate as substrate by ArgB of *Bacteroides fragilis* supports the existence in this organism of an arginine biosynthetic pathway using N-succinylated intermediates.

4.1 The context of this chapter

As stated in the introduction, in the classical arginine biosynthetic pathway, known until now, the flow of acetylated precursors leads to the formation of AOrn (Figure 2).

In the majority of organisms, AOrn is then deacetylated to form ornithine by the enzyme acetylornithine deacetylase (AOD), also called acetylornithinase (AO, E.C. 3.5.1.16), to continue the pathway.

In chapter 2 I presented results showing that in *Xanthomonads* AOrn is transcarbamylated by AOTCase producing ACit that is then deacetylated by the product of the gene *argE* to continue the arginine biosynthetic pathway in these organisms.

The discovery of the SOTCase activity of *Bf*ArgF' lead us to think that there are other possible pathways for the *de novo* arginine biosynthesis that doesn't occur through acetylated intermediates. In the case of *B. fragilis* (and concomitantly the *Bacteroidetes* family) the arginine biosynthetic pathway would involve succinylated intermediates, something never described before.

In this chapter I am going to present evidences that support this hypothesis and that let us conclude that in the *Bacteroidetes* family of microorganisms the *de novo* biosynthesis of arginine occurs through succinylated intermediates instead of acetylated intermediates, a novel pathway for the biosynthesis of arginine.

4.2 Initial attempts to assay the first enzyme of the arginine biosynthetic pathway in *B. fragilis*

Recent advances in the sequencing of the genomic DNA of several members of the family *Bacteroidetes* has made possible the localization of the distribution of the genes involved in the arginine biosynthetic pathway on these organisms (Figure 37).

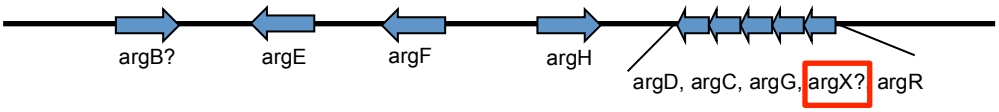


Figure 37: Schematic representation of the distribution of genes involved in the arginine biosynthetic pathway common to the *Bacteroidetes* family of microorganisms. The gene annotated as "argX?" highlighted with a red box is supposed to participate in the arginine biosynthetic pathway in these organisms. Sequence analyses using the BLASTX from NCBI returned no similarities between argX and any of the genes known to be involved in the arginine biosynthetic pathway in *E. coli*.

Located between the genes identified as *argR* and *argG* there was found an open reading frame (ORF), *argX* that codifies for a protein supposedly involved in the arginine biosynthetic pathway. Sequences analyses revealed that the protein encoded by the gene denominated *argX* does not share any homologies or similarities to any of the genes of the arginine biosynthetic pathway in other organisms.

The lack of similarities with other known genes of the arginine biosynthetic pathway could be explained by the fact that *Bacteroidetes* apparently uses an arginine biosynthetic pathway different from the rest of the organisms described until now as hypothesized before.

To test if *Bf*argX is indeed the first gene involved in the arginine biosynthetic pathway in *B. fragilis*, we cloned *argX* from *B. fragilis* TM4000 using the primers: BfargXXhoIFW and BfargXNdeIRV as described in Materials and Methods generating the expression vector pET21b-*Bf*ArgX. Should this be true the reaction catalyzed by *Bf*ArgX would be described by the following equation:

$$\text{succinyl-CoA + L-glutamate} \xrightarrow{Bf\text{ArgX}} \text{CoASH + N-succinyl-L-glutamate}$$

In an attempt to elucidate the nature of the reaction catalyzed by *Bf*ArgX, lysates of cells over-expressing the enzyme were assayed in the presence of Succinyl-CoA (SucCoA) and glutamate using the assay described by Caldovic et al.[81]. Monitorization of the product (N-succinyl-L-glutamate, NSG), formed as well as substrate consumed produced similar results. Unfortunately, the rate of product formation in the presence of the lysate was indistinguishable from that of the spontaneous reaction of the substrates (Figure 38) and consequently we could not conclude that *Bf*ArgX has N-succinylglutamate synthase (NSGS) activity.

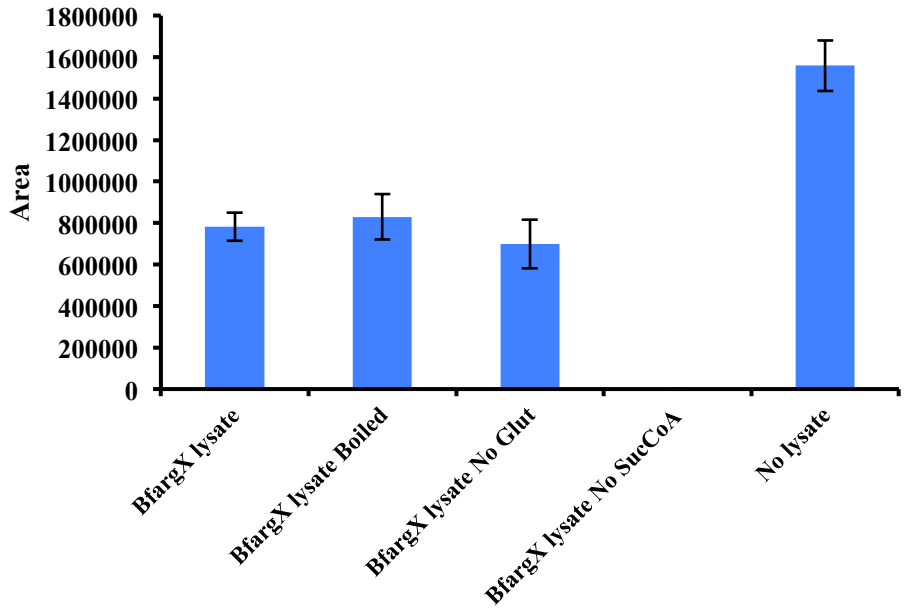


Figure 38: Initial attempts to assay the activity of *Bf*ArgX. *E. coli* BL21 (DE3) cells over expressing *Bf*ArgX were disrupted in 50mM KPi_3 pH 7.5 buffer containing 200mM KCl, 20% glycerol, 10mM β-ME, 0.0006% Triton X-100 and 1% acetone. A total of 193.6µg of total protein were assayed in 100µl of 50mM MES buffer pH 6.0 containing 50mM glutamate and 2.5mM SucCoA. The reaction was stopped by adding 100µl of 30% TCA and after centrifugation, samples were loaded onto the LC-MS instrument. Product formation was quantified using glutamate as standard and measuring the amount of substrate consumed.

4.3 Can the spontaneous formation of NSG be reduced so the activity of *Bf*ArgX could be measured?

In order to be able to assay the supposed NSGS activity of *Bf*ArgX I monitored the rate of spontaneous NSG formation from glutamate and SucCoA under different conditions to determine the conditions at which such formation is minimal and therefore more easily distinguishable from the enzymatically catalyzed reaction.

Effect of pH on the spontaneous formation of NSG

The spontaneous formation of NSG was measured at different pH using MES buffer. The conditions for the assay were the same as described in the previous section with the only difference of the pH of the solution in which the compounds were mixed.

Figure 39 shows that, under the conditions tested, the spontaneous formation of NSG from glutamate and SucCoA is maximal at pH 7.5 while minimal at pH 6.0.

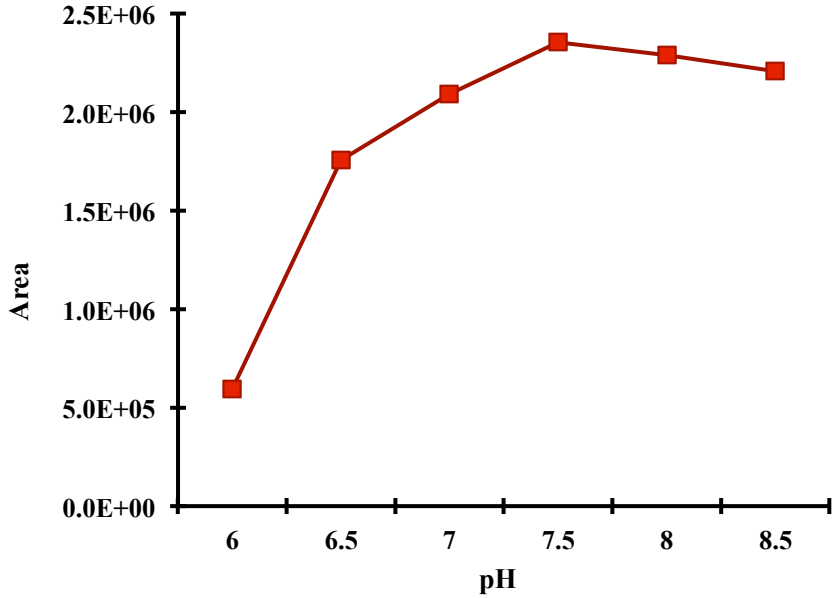


Figure 39: pH effects on the rate of spontaneous formation of NSG. The pH effects on the spontaneous formation of NSG were tested using 50mM MES buffer at different pH (ranging from 6 to 8.5). Product formation was determined by the presence of the ion whose m/z corresponded to that of NSG (248.1 m/z) on each sample measured by LC-MS.

Based on these results our first approach to elucidate the enzymatic activity of *Bf*ArgX was performed at a pH (pH 6) at which the spontaneous formation of NSG is minimal compared to higher pH. The fact that the NSG formation in the presence of the cells lysate containing *Bf*ArgX was indistinguishable from that in the absence of the lysate lead us to think that the buffer composition for the assay of this enzyme was not appropriate, therefore we tested the spontaneous formation of NSG in different buffers at pH 6.0 and 6.5.

Effect of buffer composition on the spontaneous formation of NSG at pH 6 and 6.5

The spontaneous formation of NSG in MES and Bis-Tris buffers was tested at pH 6 and 6.5 in an attempt to optimize the conditions of the assay for the enzymatic activity of *Bf*ArgX.

Figure 40 shows that the rate of spontaneous formation of NSG from glutamate and SucCoA is minimal in MES pH 6.0 compare to MES pH 6.5 and Bis-Tris pH 6.0 and 6.5.

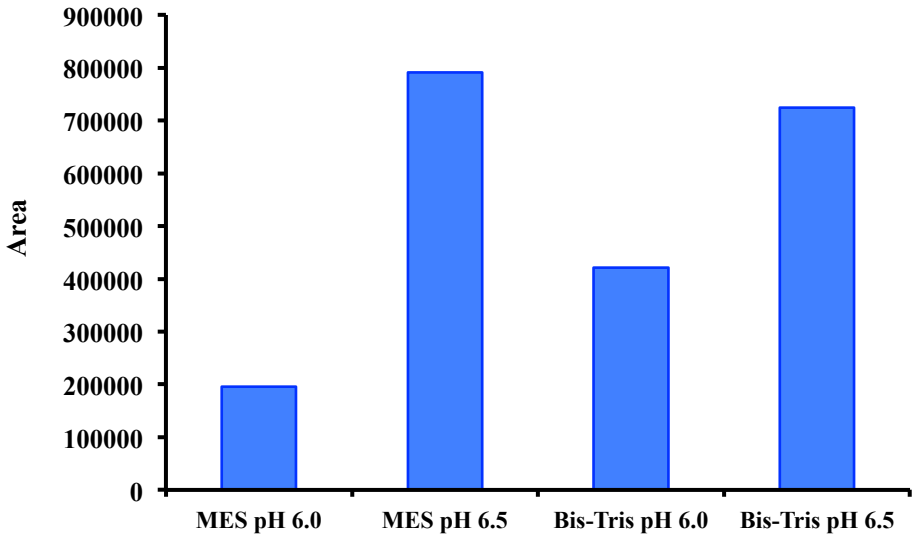


Figure 40: Determination of the rate of spontaneous formation of NSG in MES and Bis-Tris buffers at pH 6.0 and 6.5. The buffer composition effects on the spontaneous formation of NSG were tested using 50mM MES or Bis-Tris buffers at pH 6.0 and 6.5. Product formation was determined by the presence of the ion whose m/z corresponded to that of NSG (248.1 m/z) on each sample measured by LC-MS.

The results shown in Figure 40 demonstrate that the conditions in which the initial assays to elucidate the activity of *Bf*ArgX were performed were optimal in terms of minimal spontaneous formation of NSG from glutamate and SucCoA.

The lack of similarities between *Bf*argX and any known gene involved in the arginine biosynthetic pathway, and the fact that we couldn't assay the activity of the enzyme encoded by it, forced us to move to the following reaction involved in the arginine biosynthetic pathway (Figure 2): the ATP-dependent phosphorylation of NAG (in the case of *Bacteroidetes*, NSG), in order to prove that the arginine biosynthetic pathway in *B. fragilis* indeed occurs through succinylated intermediates as hypothesized before.

4.4 Synthesis of the putative substrate of the reaction catalyzed by *Bf*ArgB

The second step of the arginine biosynthetic pathway, the ATP-dependent phosphorylation of NAG (Figure 2), is catalyzed, in most organisms[9], by N-acetyl-L-glutamate kinase (ArgB, NAGK, EC 2.7.2.8).

This enzyme becomes the point of control of the metabolic flow of the arginine biosynthetic pathway in those organisms with a "cyclic pathway" (those that have an *argJ* gene). Since no other genes similar to *argJ* can be found in the *B. fragilis* genome it is expected that the control of the metabolic flow of intermediates to be exerted therefore at the level of the synthesis of NSG catalyzed by ArgA.

So far we have unsuccessfully attempted to elucidate the enzymatic activity of the enzyme encoded by the gene denominated *argX* in *B. fragilis* as shown in the

previous epigraphs and therefore we could not determine if this protein is responsible for the control of the metabolic flow of intermediates in the arginine biosynthetic pathway of *B. fragilis*.

The *B. fragilis argB* gene (Figure 35) has 45.8% similarity with its *E. coli* ortholog, and it has been identified as essential for arginine biosynthesis in this organism (Dr. Malamy, transposon mediated mutagenesis experiments, personal communication).

If the hypothesis of a succinylated initial part of the pathway is correct, the reaction catalyzed by the product of the gene *argB* in *B. fragilis*, should be the ATP-dependent phosphorylation of NSG to produce N-succinylglutamyl phosphate (NSGP) to continue the arginine biosynthetic pathway. This reaction is described by the following equation:

$$\text{N-acetyl-L-glutamate + ATP} \xrightarrow{{}_{Bf}\text{ArgB}} \text{N-acetyl-L-glutamate 5-phosphate + ADP}$$

To test this hypothesis the putative substrate of ${}_{Bf}$ArgB, NSG had to be chemically synthesized since it is not commercially available. NSG was chemically synthesized from succinic anhydride and glutamate.

0.025 moles of glutamic acid (monosodium salt) were dissolved in 25ml of water. NaOH was added until fully dissolved. The solution was then incubated with 0.025 moles of succinic anhydride for 24 hours at around 50°C with constant stirring. Quantification of the product formation was obtained by measuring the depletion of peaks corresponding to unreacted glutamic acid (148 m/z) using known concentrations of glutamic acid as a standard curve by LC-MS (Figure 41)

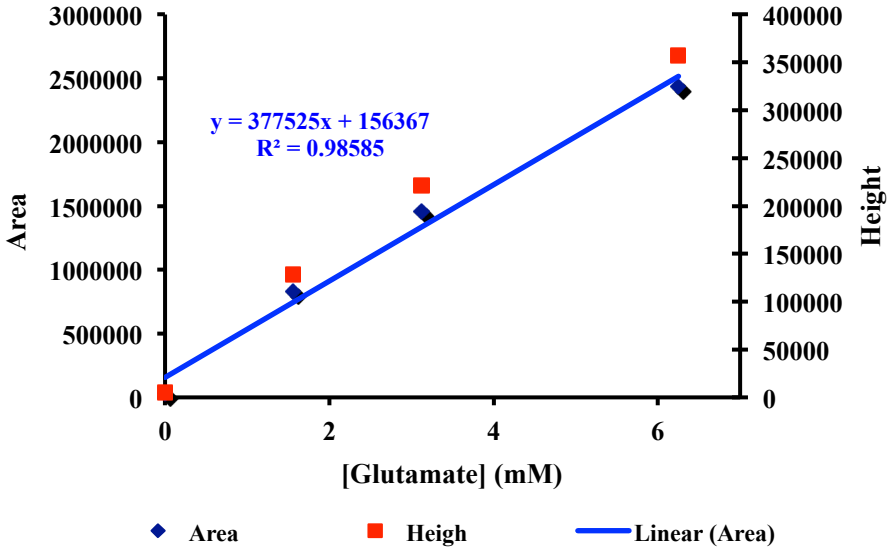


Figure 41: Glutamate standard curve used for the quantification of NSG. Chemically synthesized NSG was quantified based on the amount of unreacted glutamate present in the sample after the reaction was considered finished. LC-MS analysis tracking the ion corresponding to glutamate (148.1 m/z). Blue squares represent the area measured under the peak of the 148.1 m/z ion and red squares represent the height of the peak. Both measurements shows a good correlation but for quantification purposes I used the area results.

Both glutamate and aspartate are structurally very similar; therefore, in order to test the specificity of the enzymatic reaction catalyzed by *Bf*ArgB, assuming we could detect and quantify it, N-succinyl-L-aspartate (NSA) was also chemically synthesized from succinic anhydride and L-aspartate following the same procedure used for the synthesis of NSG.

LC-MS and ninhydrin reactivity (measured in an amino acid analyzer) techniques were used for the quantification of the product formed in each reaction, as indicated in Materials and Methods. Quantifications were done based on the amount of free, un-reacted, glutamate and aspartate present in each sample after the reactions were considered to be complete. The results were very similar using both techniques and we obtained a 99.95% and 86.8% conversion for NSG and NSA respectively yielding a concentration of 679.7 and 590.3mM respectively.

These chemically synthesized compounds were used in the enzymatic activity assays without further purification.

4.5 Demonstration that enzyme codified by *BfargB* catalyzes the ATP-dependent phosphorylation of NSG

The gene *argB* from *B. fragilis* was cloned as described in Materials and Methods using the primers BfargABNdeIFW and BfragABXhoI into the expression vector pET28a, and over-expressed in BL21 cells.

Lysates assays to test the activity of BfArgB

In a first approach to identify the activity of *Bf*ArgB, cells over expressing the protein were disrupted by sonication in Ni-column buffer A as indicated in Materials and Methods, and the lysate assayed without further purification as described by Qu et al. 2007[82].

Cells lysates were assayed for 30 minutes at 37°C in 100mM Tris buffer, pH 7.4 containing 400mM ATP, 100mM $MgCl_2$ and 400mM hydroxylamine in the presence of 8mM NAG, glutamate or aspartate or 54.4mM NSG or 52.1mM NSA. H_2O and boiled cells lysate were used as negative controls of the reaction (Figure 42). Product formation was quantified using glutamyl hydroxamate as standard.

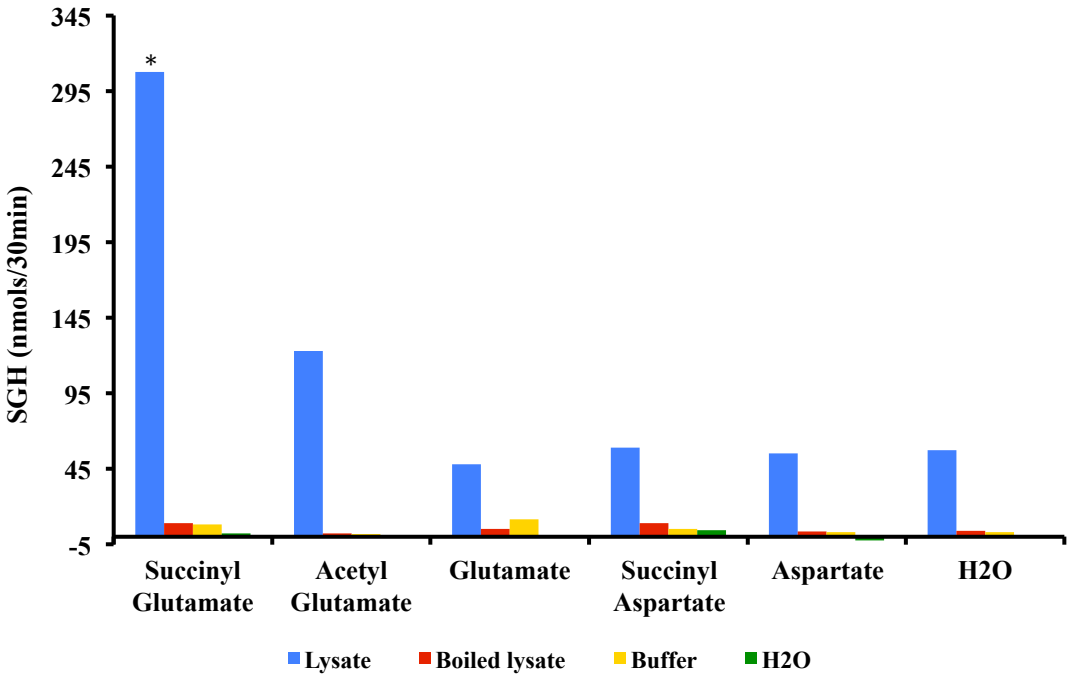


Figure 42: Initial attempt to assay the activity of *Bf*ArgB using chemically synthesized NSG. Cells over-expressing *Bf*ArgB were lysated in Ni-column buffer A (Materials and Methods) and assayed for kinase activity without further purification. Each sample contained 610.2µg of total lysate (boiled or non boiled) in 100mM Tris buffer, pH 7.4 containing 400mM ATP, 100mM $MgCl_2$ and 400mM hydroxylamine and 8mM NAG, glutamate or aspartate or 54.4mM NSG or 52.1mM NSA in a final volume of 500µl. After 30 minutes at 37°C the reaction was terminated by addition of 500µl of ferric chloride solution (5% $FeCl_3$, 8% TCA and 0.3 M HCl). Product formation was estimated using glutamyl hydroxamate as standard. (*****) indicates that the sample absorbance was over the limit of detection of the spectrophotometer used and therefore should only be considered as a reference for the presence of the kinase activity on the cells lysate and not as an accurate quantification of such activity.

The results obtained in this experiment showed for the first time that *Bf*ArgB has N-succinylglutamate kinase (NSGK) activity.

To confirm that *Bf*ArgB has NSGK activity and that the results presented in the previous figure were not an artifact due to the complexity of the sample assayed (cells over-expressing *Bf*ArgB lysate), the assay was performed with the purified enzyme.

Purified BfArgB activity confirmed that this enzyme is in fact a N-succinylglutamate kinase

*Bf*ArgB was purified using Ni affinity as described in Materials and Methods with purity higher than 95%.

To test the specificity of the reaction catalyzed by *Bf*ArgB, 10µg of purified enzyme were assayed in 100mM Tris buffer pH 6.0 containing 400mM ATP, 100mM $MgCl_2$ and 400mM hydroxylamine and 8mM NAG, glutamate or aspartate or 54.4mM NSG or 52.1 NSA in a final volume of 100µl. After 20 minutes incubation at

37°C the reaction was stopped with 100µl of ferric chloride solution (Materials and Methods). Product formation was quantified using glutamyl hydroxamate as standard. Heat-inactivated enzyme (10µg of purified *Bf*ArgB boiled for 10 minutes) was used as negative control for the experiment (Figure 43).

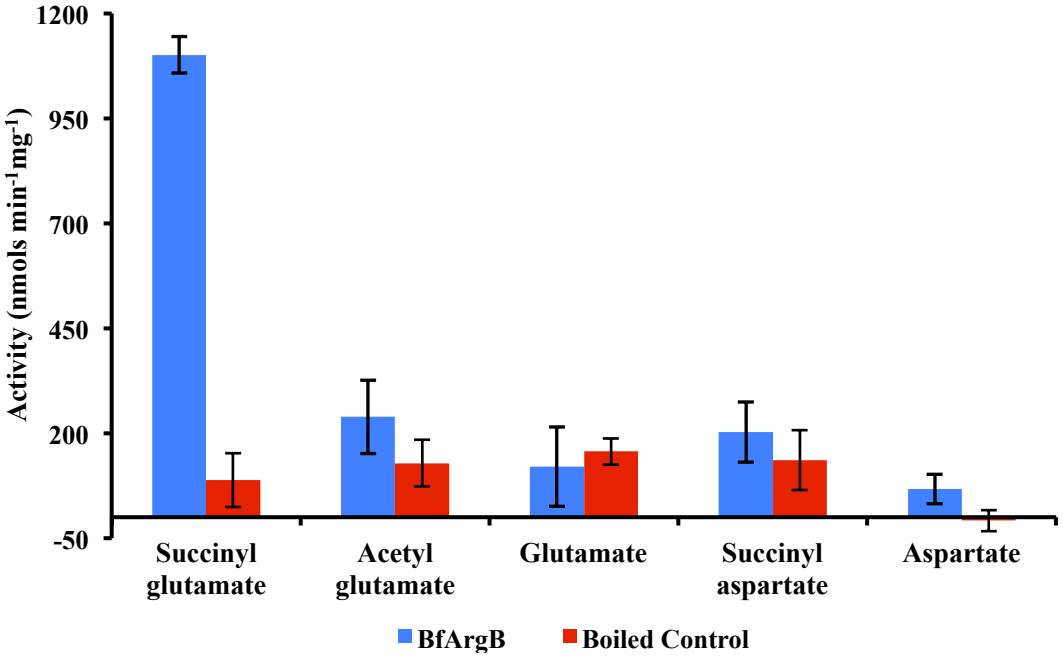


Figure 43: Confirmation that *Bf*ArgB catalyzes the ATP-dependent phosphorylation of succinylglutamate. 10µg of purified *Bf*ArgB were assayed for NSGK activity in 100mM Tris buffer pH 7.4 containing 400mM ATP, 100mM $MgCl_2$ and 400mM hydroxylamine and 8mM NAG, glutamate or aspartate or 54.4mM NSG or 52.1 NSA in a final volume of 100µl. The reaction was carried out for 20 minutes at 37°C and stopped with 100µl of ferric chloride solution (Materials and Methods). Each assay was performed in triplicate and the graph shows the average and standard deviation for each condition.

These results prove that *Bf*ArgB has NSGK activity and highlight its specificity, being able to catalyze only the ATP-dependent phosphorylation of NSG but not of any of the other substrates tested (NAG, glutamate, NAS or aspartate).

These results confirmed the initial hypothesis of an arginine biosynthetic pathway in *B. fragilis* that utilizes succinylated instead of acetylated intermediates, which has never been described before.

4.6 Initial characterization of the NSGK activity of *Bf*ArgB

As in the case of the enzymes described previously in this thesis work, I did an initial characterization of the reaction catalyzed by *Bf*ArgB and studied the putative inhibitory effects of arginine to elucidate if *Bf*ArgB controls the arginine biosynthetic pathway flow.

Test for enzyme inactivation and linearity of the reaction catalyzed by BfArgB

The linearity of the reaction catalyzed by BfArgB was tested using 2 and 2.5µg of purified enzyme. This test was also used to determine if enzyme inactivation occurs during the time of the reaction. BfArgB was assayed in 100mM Tris buffer pH 6.0 containing 20mM ATP, 100mM $MgCl_2$ and 400mM hydroxylamine and 8mM NSG for 1, 2.5, 5, 7.5 and 10 minutes at 37°C.

The results showed that the reaction catalyzed by BfArgB is linear with respect to the time and that there is no detectable enzyme inactivation in the course of the reaction when either 2 or 2.5µg of purified BfArgB are used for the assay (Figure 44).

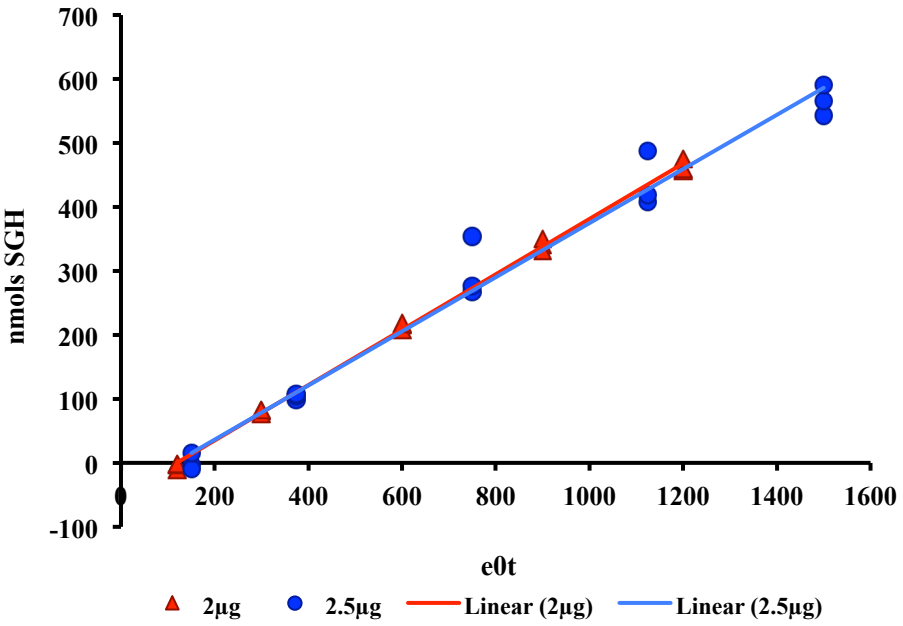


Figure 44: Test for enzyme inactivation (Selwyn's test) and linearity of the reaction catalyzed by BfArgB. 2 and 2.5µg of purified BfArgB were assayed in 100mM Tris buffer pH 6.0 containing 20mM ATP, 100mM $MgCl_2$ and 400mM hydroxylamine and 8mM NSG for 1, 2.5, 5, 7.5 and 10 minutes at 37°C. In this graph the results are presented as nmols of succinylglutamate hydroxymate against e0t (time of reaction x enzyme concentration). These results prove that the reaction catalyzed by BfArgB is linear and that there is no detectable enzyme inactivation when either 2 or 2.5µg are used for the assay.

Based on these results we decided to use 2.5µg of enzyme for the characterization of the activity of BfArgB.

Determination of the kinetics parameters K_m and V_{max} for NSG of BfArgB

To determine the K_m and V_{max} for NSG of BfArgB, 2.5µg of purified enzyme were assayed in 100mM Tris buffer pH 6.0 containing 100mM $MgCl_2$, 400mM

hydroxylamine in the presence of a constant ATP concentration of 20mM while varying the NSG concentration from 0 to 20mM (Figure 45). After 10 minutes incubation at 37°C the reaction was stopped by adding fresh ferric chloride solution (Materials and Methods).

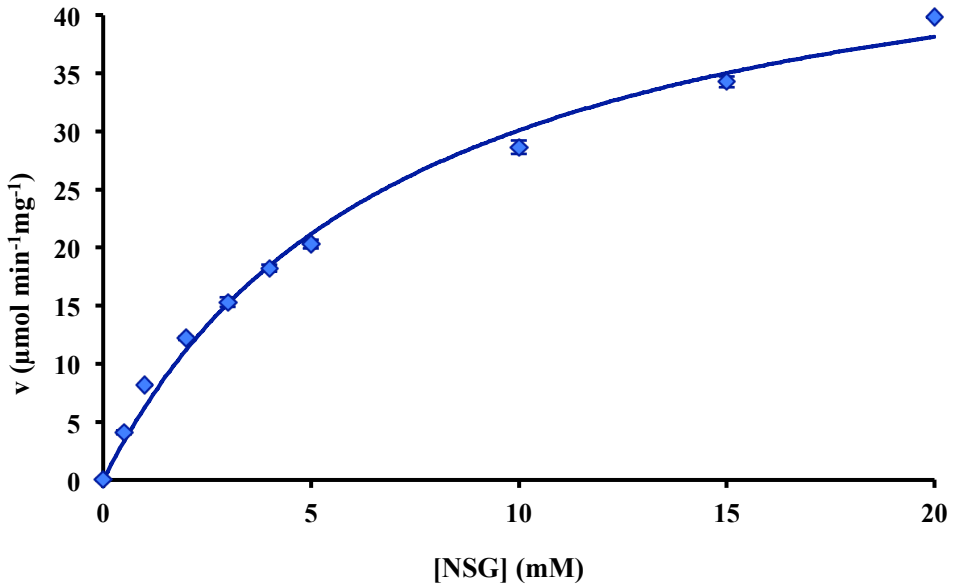


Figure 45: Determination of the kinetics parameters for NSG of *Bf*ArgB. 2.5µg of purified *Bf*ArgB were assayed in 100µl of 100mM Tris buffer pH 6.0 containing 100mM $MgCl_2$, 400mM hydroxylamine in the presence of a constant ATP concentration of 20mM while varying the NSG concentration from 0 to 20mM. The reaction was carried out at 37°C and after 10 minutes stopped by the addition of 100µl of fresh ferric chloride solution. Data were fitted to the Michaelis-Menten equation for enzyme activity using the program GNUPLOT. Each data represents the average and standard deviation of 3 independent assays per condition. Product formation was quantified using glutamyl hydroxamate as standard.

As Figure 45 shows the NSG saturation curve for *Bf*ArgB follows a hyperbolic of Michaelis-Menten kinetic. Experimental data were fitted to the corresponding equation using non-linear fitting using the program GNUPLOT. The kinetic parameters determined for NSG in this experiment were:

K_m: 7.3 ± 0.44mM and
V_{max}: 52.1 ± 1.4µmol min^{-1} mg^{-1}

The K_m for NSG of *Bf*NSGK is much higher with respect to other NAGKs for NAG (200 µM for the *E. coli* enzyme[83]), only *P. aeruginosa* NAGK has a similar K_m being only 3 times lower[84].

*Determination of the kinetic parameters K_m and V_{max} for ATP of Bf*ArgB

For the experimental determination of the kinetic parameters K_m and V_{max} for ATP of *Bf*ArgB we used 2.5µg of purified enzyme as described in the previous

epigraph. In this experiment the concentration of NSG was kept constant at 40mM while the ATP concentration was varied from 0 to 20mM. The Mg^{2+} concentration was kept 10mM higher than that of the ATP at each condition assayed.

The ATP saturation curve follows a Michaelis-Menten kinetic as shown in Figure 46. And the calculated kinetic parameters are:

K_m: 4.6 ± 0.25mM and
V_{max}: 54.7 ± 1.1 µmol min^{-1} mg^{-1}

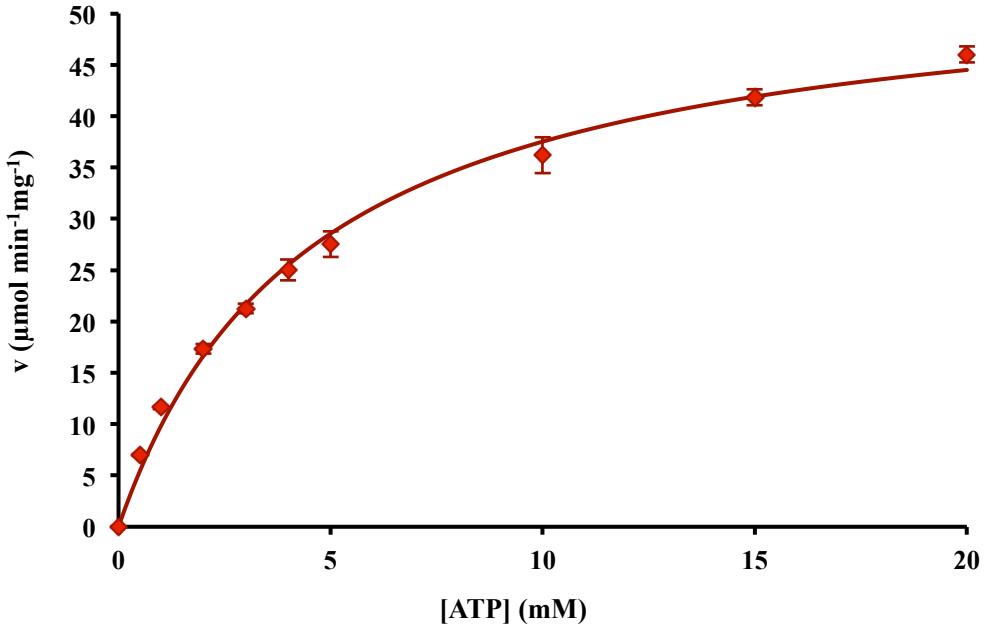


Figure 46: Determination of the kinetic parameters K_m and V_{max} for ATP of *Bf*ArgB. 2.5µg of purified *Bf*ArgB were assayed at 37°C in 100µl of 100mM Tris buffer pH 6.0 containing 400mM hydroxylamine in the presence of a constant NSG concentration of 40mM while varying the ATP concentration from 0 to 20mM. Mg^{2+} concentration was kept 10mM higher than the ATP concentration throughout the experiment. After 10 minutes the reaction was stopped by the addition of 100µl of fresh ferric chloride solution (Materials and Methods). Data were fitted to the Michaelis-Menten equation for enzyme kinetic using non-linear fitting with the software GNUPLOT.

In this case the K_m for ATP is also higher than that of the *E. coli* enzyme (300 µM[83]), but not that different from *P. aeruginosa* NAGK (3 mM[84]).

The K_m calculated for both substrates of *Bf*ArgB are among the highest compared to other ArgB enzymes from organisms that posses an acetylated pathway for arginine biosynthesis. It is possible that *Bf*ArgB forms a complex with ArgA' or N-succinylglutamate synthase or the enzyme responsible for NSG synthesis to compensate for the low affinity of the substrates. It is also possible that *Bf*ArgB interacts with other proteins in order to be more effective.

Since the activity of the enzyme that we thought could be the first enzyme of the arginine biosynthesis pathway in *B. fragilis* (ArgX) could not be assayed it made

impossible to investigate the possible regulatory effects that arginine may have on it. Therefore we decided to study the effects that arginine may have on the activity of *Bf*ArgB to determine whether or not this enzyme could be a point of control in the arginine biosynthetic pathway in the *Bacteroidetes* family of microorganisms.

Study of the arginine effect on the activity of BfArgB

To study the possible effects of arginine on the activity of *Bf*ArgB we assayed 2.5µg of purified enzyme using the assay described before (Materials and Methods) at saturating conditions (40mM NSG, 40mM ATP), K_m conditions (NSG 10mM, ATP 6mM) and half K_m conditions (NSG 3.6mM, ATP 2.3mM) in the presence of several arginine concentrations ranging from 0 to 50mM (Figure 47). The Mg^{2+} concentration was kept 10mM higher than that of ATP at all times.

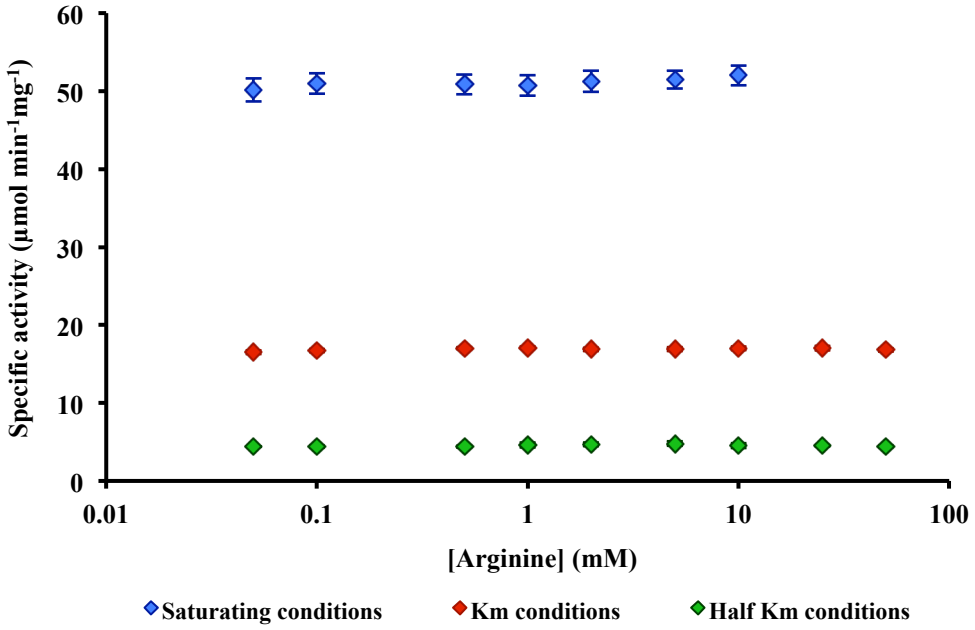


Figure 47: Study of the effects of arginine on *Bf*ArgB activity. 2.5µg of purified *Bf*ArgB were assayed in 100µl of 100mM Tris buffer pH 6.0 containing 400mM hydroxylamine in the presence of several arginine concentrations ranging from 0 to 50mM at saturating conditions (40mM NSG and 40mM ATP), Km conditions (10mM NSG and 6mM ATP) and half Km conditions (3.6mM NSG and 2.3mM ATP) for 10 minutes at 37°C. Mg^{2+} concentration was kept 10mM higher than the ATP concentration. Product formation was quantified using glutamyl hydroxamate as standard. Logarithmic scale was used only for representation purposes only.

The results obtained in these assays shown in Figure 47 demonstrate that arginine has no effect whatsoever on the activity of *Bf*ArgB.

The lack of inhibitory effects of arginine on the activity of *Bf*ArgB suggest that in the organisms which posses a succinylated arginine biosynthetic pathway the metabolic flow of intermediates is controlled at the level of NSG synthesis, which also suggest a linear non-classic pathway rather than a cyclic pathway. This

hypothesis is supported also by the fact that in the genome of the members of the *Bacteroidetes* family the gene that codifies for an enzyme similar to OAT could not be found.

4.6 Summary

In this Chapter 5: **"The use of N-succinyl-L-glutamate as substrate by ArgB of *Bacteroides fragilis* supports the existence in this organism of an arginine biosynthetic pathway using N-succinylated intermediates"**, I have presented further evidences that support the hypothesis, developed from the discovery of the SOTCase activity of the *argF'*-encoded protein in *B. fragilis*, of the existence of an arginine biosynthetic pathway in the *Bacteroidetes* family of microorganisms that uses N-succinylated instead of N-acetylated intermediates.

Unfortunately the activity of the, hypothesized, first enzyme of the pathway (ArgX, Figure 35) could not be assayed due to the high rate of spontaneous succinylation of L-glutamate in the presence of SucCoA. It is possible that the enzyme requires the interaction with other members of the arginine biosynthesis pathway (like ArgB), or the presence of certain cofactors in order to catalyze the first reaction of the pathway. Determination of the means by which N-succinylglutamate enters the arginine biosynthesis pathway in these organisms could provide relevant information about the regulation and control of the metabolic flow in this pathway.

Although I could not assay the very first reaction involved in the arginine biosynthetic pathway in *B. fragilis*, I was able to demonstrate that the *argB'*-encoded protein from this organism catalyzes the phosphorylation of NSG, probing that the biosynthesis or arginine is mediated by succinylated intermediates as hypothesized after the discovery of the SOTCase activity of *Bf*ArgF'.

*Bf*ArgB has been partially characterized as a result of the present thesis work and the kinetic parameters K_m and V_{max} determined for the two substrates of the enzyme, NSG and ATP (Table 5).

Table 5: Summary of the kinetic parameters determined for *Bf*ArgB.

	Km (mM)	Vmax (μmols min^{-1} mg^{-1})
NSG	7.3 ± 0.44	52.1 ± 1.4
ATP	4.6 ± 0.25	54.7 ± 1.1

The affinity of the enzyme for the two substrates is low compared to NAGSs from other microorganisms that use acetylated intermediates[83, 84]. Proximity of the substrates would help to compensate for the low affinity of *Bf*ArgB for them, therefore it is possible that ArgB and ArgX in *B. fragilis* interact with each other in such a way that ArgX becomes active and the substrates are easily available to ArgB to continue the arginine biosynthesis pathway as is the case of the *S. cerevisiae* and *N. crassa* enzymes[17].

The absence of an OAT-like gene in the genome of *B. fragilis* led us to hypothesize that the arginine biosynthetic pathway in this organism is linear rather

than cyclic. Further confirmation of the linear nature of the arginine biosynthetic pathway in *B. fragilis* was provided by the study of the arginine effects on *Bf*ArgB activity. As shown in Figure 45, arginine has no effects on the activity of *Bf*ArgB probing that this reaction is not regulated by the final product of the pathway and therefore the arginine biosynthetic pathway in these organisms is most likely to be linear instead of cyclic.

Resumen en Español

Resumen de la tesis

El amino ácido L-arginina es un componente esencial de todos los organismos vivos. Su importancia reside en la gran variedad the funciones que la arginina, así como los intermediarios involucrados en su síntesis, tienen en las células, como muestra la Figura 1.

Entre las principales funciones de la arginina en la célula se encuentran:

- forma parte de la composición amino acídica de muchas proteínas
- precursor de los siguientes compuestos usados por las células para almacenar energía
    - fosfato de arginina, usado en la contracción muscular en invertebrados y
    - fosfato de creatina, que se encuentra almacenado en el músculo y el cerebro en vertebrados
- metabolito intermediario en el ciclo de la urea
- precursor de las moléculas de señalización: oxido nítrico, agmatina y glutamato en mamíferos
- y es usada como compuesto de almacenamiento de nitrógeno en plantas y cianobacterias

En numerosos procariotas, hongos y plantas la síntesis *de novo* de la arginina ocurre a partir the glutamato en ocho pasos enzimáticos como se describe en la Figura 2. La acetilación de los precursores iniciales en la biosíntesis de arginina permite la separación de ésta ruta metabólica de la biosíntesis de prolina.

De todos los pasos implicados en la ruta biosintética de la arginina, la acetilación de glutamato y su fosforilación han sido los más estudiados dada su importancia en la regulación del flujo de metabolitos en ésta ruta.

Sin embargo, la transcarbamilación de la ornitina constituye otra reacción crucial en la ruta de biosíntesis y biodegradación de la arginina como muestra la Figura 4.

En la ruta de biosíntesis, ornitina es transcarbamilada por la enzima ornitina transcarbamilasa (OTC, EC 3.2.1.1). OTC pertenece a la familia de las transcarbamilasas cuyos miembros son identificados fácilmente basándose únicamente en las homologías a nivel de las secuencias peptídicas que presentan todos ellos ya que todos comparten el motivo de unión al fosfato de carbamilo, que es un sustrato común a todos los miembros de la familia. Los miembros más estudiados de la familia de las transcarbamilasas son OTC y aspartato transcarbamilasa (ATC).

Los estudios sobre la evolución de las transcarbamilasas realizados por Labedan y colaboradores revelaron la presencia de ésta actividad en el último ancestro común universal (Last Universal Common Ancestor, LUCA), destacando la importancia de la reacciones catalizadas por las transcarbamilasas.

Más recientemente, Naumoff y colaboradores, entre otros, han puesto de manifiesto cuán inexacta es la asignación de ciertas actividades enzimáticas basándose exclusivamente en las similitudes entre nuevas secuencias genómicas y las estudiadas bioquímicamente previamente. El trabajo de Naumoff y colaboradores se centró en el caso particular de las transcarbamilasas de putrescina (PTC), miembros de la familia de las transcarbamilasas erróneamente anotadas como OTC.

La topología del árbol evolutivo generado por Naumoff y colaboradores usando las 245 secuencias disponibles en el momento (Figura 5) no difiere de la del árbol presentado por Labedan y colaboradores anteriormente. En su análisis, Naumoff y colaboradores introdujeron dos nuevos grupos de transcarbamilasas (PTC and UTC, o transcarbamilasas de actividad desconocida). El grupo de las UTC está formado por las transcarbamilasas anotadas como OTCasas pero que carecen de la capacidad para usar ornitina como sustrato, como es el caso de la enzima, potencialmente transcarbamilasa, presente en el organismo *Bacteroides fragilis*.

El Dr. Malamy, usando mutagenesis mediada por transposon, demostró que en *B. fragilis* el gen que codifica para la putativa ornithina transcarbamilasa está directamente involucrado en la biosíntesis de arginina dado que cuando dicho gen no estaba presente en el genoma, el organismo se volvía auxotrofo para la arginina.

El gen *$_{Bf}$argF´* fue clonado y la proteína purificada y cristalizada por el Dr. Shi (Figura 7). La estructura tridimensional de la enzima presenta grandes similitudes con las estructuras de OTC y ATC, desafortunadamente la actividad de esta enzima no pudo ser identificada dado que no cataliza la transcarbamilación de la ornitina.

Usando como referencia la secuencia aminoacídica de la proteína de *B. fragilis* se pudieron identificar proteínas similares a ésta en los genomas de varios organismos entre los que se encuentran *Xylella fastidiosa*, *Cytophaga hutchinsonii*, *Xanthomonas campestris*, *Prevotella ruminicola* y *Tannerella forsythus*. Todas estas proteínas identificadas compartían con la de *B. fragilis* la ausencia del motivo de unión a ornitina: SMG, presente en todas las OTCasas. Asímismo, ninguno de los genes identificados en estos organismos pudo complementar los mutantes de *E. coli* ($argF^-$, $argI^-$) que carecen de OTCasa.

La primera actividad que se descubrió de entre las anteriormente mencionadas transcarbamilasas de actividad desconocida fue la de *X. campestris* ArgF´. Dr. Morizono descubrió que la actividad de *$_{Xc}$*ArgF´ era la de una nueva transcarbamilasa: acetilornitina transcarbamilasa (AOTCasa).

La presente tesis se centra en la caracterización bioquímica de la actividad de la enzima *$_{Xc}$*ArgF´ así como en la identificación y caracterización de la actividad catalítica de la enzima *$_{Bf}$*ArgF´ y en las implicaciones que la presencia de estas dos enzimas tienen para la ruta de biosíntesis de la arginina en los organismos que las poseen.

A continuación presento el sumario de los resultados obtenidos durante la realización del presente proyecto de tesis: “Discovery of novel pathways of microbial arginine biosynthesis”.

Capitulo 2: "Functional characterization of a novel type of transcarbmylase that uses acetylornithine, identified in *Xanthomonas campestris pv campestris*"

En este capitulo se han presentado pruebas concluyentes demostrando que la actividad catalizada por la enzima $_{Xc}$ArgF´ es la de una acetilornitina transcarbamilasa (AOTCasa) mediante la identificación tanto de los sustratos como de los productos de la reacción.

La especificidad de la reacción catalizada por la $_{Xc}$AOTCasa ha sido demostrada indicando que acetilornitina es el único compuesto que junto con carbamilfosfato puede ser transcarbamilado para formar acetilcitrulina. Asimismo se han optimizado las condiciones para el ensayo de esta nueva actividad AOTCasa.

Los parámetros cinéticos (K_m y V_{max}) han sido determinados para ambos sustratos (AOrn y CP, Tabla 3).

Al contrario que ocurre con otros miembros de la familia de las transcarbamilasas, $_{Xc}$AOTCasa no es inhibida por el sustrato de la reacción (AOrn), incluso a concentraciones tan elevadas como 50mM (Figura 21). Estos resultados parecen indicar que la reacción catalizada por $_{Xc}$AOTCasa no es crucial en el control de la ruta biosintética de la arginina, sin embargo el hecho de que la ausencia de este gen confiera auxotrofía para la arginina hace resaltar su importancia en la mencionada ruta de biosíntesis.

Asimismo en este capitulo se ha demostrado que la acetilación del grupo alfa-amino no sólo es necesaria para el reconocimiento del sustrato sino que también es un requerimiento para los inhibidores, como muestran los resultados obtenidos en los estudios de inhibición por norvalina y acetilnorvalina presentados en las figuras 21 y 22, así como en los estudios con PALO y PALAO presentados en las figuras 25 y 26.

Capitulo 3: "Demonstration that a structurally obvious transcarbamylase from *Bacteroides fragilis*, of previously unknown activity catalyzaes the transcarbamylation of succinylornithine. Funtional characterization of this transcarbamylase"

En este capitulo se han presentado los resultados que prueban que la actividad enzimática de la proteína codificada por el gen *argF´* de *B. fragilis* es la de una succinilornitina transcarbamilasa (SOTCasa).

Los análisis por NMR y LC-MS confirmaron que el producto de la reacción catalizada por $_{Bf}$SOTCasa en presencia de carbamilfosfato y succinilornitina es succinilcitrullina.

Así mismo se han presentado los resultados de la caracterización inicial de la enzima $_{Bf}$SOTCasa y se han determinado los parámetros cinéticos para ambos sustratos de la reacción (Tabla 4).

Al contrario que la $_{Xc}$AOTCasa, $_{Bf}$SOTCasa muestra inhibición por sustrato (Figura 33) aunque dicha inhibición es mucho menos pronunciada comparada con otros miembros de la familia de las transcarbamilasas.

Al igual que ocurría con *Xc*AOTCasa, la *Bf*SOTCasa requiere la presencia del grupo succinilo unido al extremo alfa-amino tanto para el reconocimiento del sustrato (succinilornitina), como de los inhibidores. Esto ha sido demostrado mediante estudios de inhibición por PALO y PALSO (sintetizado químicamente).

El descubrimiento de la actividad SOTCasa de la enzima *Bf*ArgF´ unido al echo de que la ausencia del gen *argF´* en *B. fragilis* transforma este organismo en auxotrofo para la arginina indica que tanto succinilornitina como succinilcitrulina son intermediarios en la ruta de biosíntesis de la arginina en *Bacteroidetes*.

Lo expuesto en el anterior párrafo implica que la parte inicial de la ruta de biosíntesis de la arginina en los organismos que poseen actividad SOTCasa transcurre a través de intermediarios succinilados y no acetilados como ocurre en el resto de los organismos descritos hasta el momento.

Esta hipótesis es el punto de partida para los resultados expuestos en el siguiente capitulo.

Capitulo 4: "The use of N-succinyl-Lglutamate as substrate by ArgB of *Bacteroides fragilis* supports the existence in this organism o fan arginine biosynthetic pathway using N-succinylated intermediates"

En este capitulo se presentan los datos obtenidos en el estudio de la actividad enzimática de los dos primeros genes supuestamente involucrados en al biosíntesis de la arginina en *B. fragilis* (*argX* and *argB*, Figura 37).

Ambos genes fueron clonados y las respectivas proteínas expresadas y purificadas, desafortunadamente la actividad enzimática de la proteína que hipotéticamente catalizaría el primer paso de la ruta, la succinilización del glutamato, no pudo ser ensayada dado que la velocidad de la reacción espontánea era similar o superior a la de la reacción en presencia de la proteína codificada por el gen *Bf*argX (Figura 38). Posiblemente la enzima necesite unas condiciones diferentes a las ensayadas para ser activa, como por ejemplo algún cofactor desconocido de momento, o la interacción con otras proteínas.

Por consiguiente los esfuerzos se concentraron en el ensayo de la posible actividad succinilglutamato kinasa de la proteína codificada por el gen *argB* de *B. fragilis*. Para poder ensayar esta actividad, el potencial sustrato succinilglutamato, fue sintetizado químicamente y dada el alto porcentaje de conversión (Figura 41), usado sin posterior purificación en los ensayos enzimáticos.

Como se muestra en las figuras 42 y 43, sólo el succinilglutamato puede ser utilizado como sustrato por la enzima *Bf*ArgB, lo cual demuestra que la biosíntesis de la arginina en *B. fragilis* transcurre a través de intermediarios succinilados y no acetilados tal y como se había hipotetizado tras el descubrimiento de la actividad SOTCasa de *Bf*ArgF´.

Usando el succinilglutamato sintetizado químicamente en el laboratorio, la actividad NSGK de *Bf*ArgB ha sido caracterizada y se han calculado los parámetros cinéticos (K_m y V_{max}, Tabla 5).

A continuación procederé a la discusión de los resultados presentados en los anteriores capítulos de la presente tesis.

Chapter 5. General discussion: Novel pathways of arginine biosynthesis.

The study of the arginine biosynthetic pathway has mainly focused, for the last years, in the regulation and control of the flow of its metabolites. As a consequence new genes that codifies for different enzymes that are able to initiate and/or regulate the arginine biosynthetic pathway have been discovered (shorter NAGS of the Gcn5-related N-acetyltransferase (GNAT) family, either independent or fused with the ArgH protein (argininosuccinase), bifunctional ornithine acetyltransferase (OAT), that is able to acetylate glutamate with both AcCoA and AOrn[1]).

Due to the focus on the initial component of the arginine biosynthetic pathway other aspects of this pathway have been neglected, in part also because of the belief that the biosynthesis of arginine uses a canonical pathway either linear (or classic) or cyclic (Figure 2).

Recent advantages in the sequencing of the genome of an increasing number of organisms have revealed the existence of, among many other genes, different versions of OTCase. Nevertheless these versions have been erroneously annotated as OTCases based on pairwise sequence similarities at the level of the primary structure, as the work of Naumoff et al.[46] and Sekowska et al.[48], among others[49-51] highlighted.

The work leading to the present Thesis started with the study of the arginine biosynthetic pathway in the anaerobe *B. fragilis* with the objective of creating an organism able to overproduce arginine for its use in urea cycle and other related disorders patients to treat hyperammonemia. While studying the arginine biosynthesis in *B. fragilis*, Dr. Malamy identified a gene (*Bf*argF) that codified for an OTC-like protein that was essential for the synthesis of arginine in this organism. Although the crystal structure of the *Bf*ArgF' protein was solved showing a high degree of similarity with that of OTC from other organisms, the enzymatic activity could not be determined and remained a mystery until recently[85].

As stated in the introduction to Chapter 2, the genomic sequence of *Bf*argF' was used to identify several other genes in different organisms in which the sequence signatures of OTC were modified (Figures 8 and 9). Among these newly identified genes were those of *Xanthomonas campestris pv campestris* and *Cytophaga hutchinsonii* which played a key role in the identification of the enzymatic activity of some of these OTC-like proteins and that constituted the basis over which the present Thesis work was developed.

We have discovered two new members of the transcarbamylases family of enzymes: N-acetylornithine and N-succinylornithine transcarbamylase. These two new transcarbamylases had expanded our view of the, until now, "canonical" *argF* gene involved in the arginine biosynthetic pathway. The objective of this project was to describe these two members of the transcarbamylases family and study the arginine biosynthetic pathway in the organisms that contains them to first, extend our knowledge of the pathway and second, provide some of the basis for developing new strategies against the pathogenic effects of such microorganisms.

Characterization of the AOTCase activity of *Xc*ArgF'

Before the present thesis project started, acetylornithine had been identified as the substrate of *Xc*ArgF' by Dr. Morizono.

Dr. Morizono's experiments proved that *Xc*ArgF' was able to catalyze the transcarbamylation of acetylornithine to produce a compound with an ureido group on its formulation, most likely acetylcitrulline.

I studied the chemical nature of the products of the reaction catalyzed by *Xc*ArgF' as well as characterized its kinetic behavior in the presence of both substrates (CP and AOrn) and inhibitors. I have also synthesized and assayed the effects of two potential inhibitors of *Xc*ArgF': acetylnorvaline and PALAO.

The results obtained using an amino acid analyzer and LC-MS (Figures 11 and 12 respectively) indicated that the product of the reaction catalyzed by *Xc*ArgF' contained an ureido group that interferes with the free amine reaction with ninhydrin (Figure 11) and that the retention time and m/z matches that of the purified chemically synthesized ACit (Figure 12). Further confirmation that the product of the reaction was ACit was obtained from the NMR analyses performed using COSY-NMR (Figure 13). All these experiments along with the identification of acetylornithine as the substrate of *Xc*ArgF' allowed us to catalog this enzyme as a new member of the transcarbamylases family: N-acetyl-L-ornithine transcarbamylase (AOTC).

Once that *Xc*ArgF' was proven to be an AOTCase I proceeded to characterize its activity. I was able to isolate *Xc*AOTCase with a high degree of purity (Figure 14) in order to use it for enzyme activity assays.

The transcarbamylation reaction catalyzed by *Xc*AOTC is highly specific for acetylornithine as Figure 15 shows and follows a Michaelis-Menten type of kinetic for both substrates, AOrn and CP (Figures 18 and 19) described by the kinetic parameters summarized in Table 3. The affinity of *Xc*AOTCase for its substrates is among the highest compared to other members of the family[41, 68-71], which could be explained by the differences in the orientation of the binding of AOrn compared to the orientation of PALO or norvaline to OTC[86]. In *Xc*AOTC the carboxyl group of AOrn is rotated approximately 120° between the Cγ-Cδ bond relative to the carboxy group of ornithine in OTC structures. The slightly higher energy of this conformation is probably one of the reasons for the lower affinity of AOrn to *Xc*AOTC compared to Orn for OTC[86].

*Xc*AOTC shows only slight substrate inhibition when AOrn concentrations were as high as 50mM (Figure 20), with a calculated K_{si} of 231mM, while in the case of human OTCase this value is 2.5mM[41]. The lack of substrate inhibition can be explained by the random binding mechanism of the reaction of *Xc*AOTC[86] compared to the ordered mechanism of the reactions catalyzed by ATC and OTCases[52, 69, 87, 88]. The high K_{si} for AOrn could be explained by the fact that the AOrn binding pocket seems to be preformed before and not upon AOrn binding[86] in contrast to ATC and OTCases in which the second-substrate binding pocket is an induced-fit pocket, where binding of the second substrate induces a conformational change and brings the flexible 240's loop in the active site to interact with the second substrate[89, 90]. In *Xc*AOTC movement of the 240's loop and domain closure seem to be restrained by the stronger domain-domain interactions provided by the extended 120's and 240's loops[86].

The high specificity of *Xc*AOTC for AOrn led us to hypothesize that αN-acetylation is necessary not only for substrate recognition but for inhibitors to bind the active site of the enzyme. To corroborate this hypothesis I tested the acetylated versions of two known OTC inhibitors: norvaline and PALO. The results showed that ANorv behaves as a competitive inhibitor of *Xc*AOTCase (Figure 22), while Norv had a very small effect on the V_{max} at a much higher concentration (Figure 21). The calculated K_i of ANorv for *Xc*AOTC (0.3mM) is much higher than that of Norv for OTC (5.8μM)[72] which could be explained by different interactions of the inhibitors to the respective enzyme since AOrn binds to *Xc*AOTC in a different orientation from those observed in PALO or Norv complexed OTCases[86, 90-92].

PALAO was proven a much more efficient inhibitor of *Xc*AOTC with a calculated K_i of 22.4μM (Figures 24 and 25), although, compared to PALO is ~23 times less effective (K_i of PALO for OTC is 0.77μM[75]). The fact that PALO did not inhibit the activity of *Xc*AOTC proved one more time that αN-acetylation is essential for both substrate and inhibitors recognition.

Other potential *Xc*AOTC inhibitor tested was pyrophosphate. Pyrophosphate showed significant effects over the activity of *Xc*AOTC reducing the V_{max} by 40% at 5mM (Figure 23), unfortunately the inhibitory mechanism could not be determined based on the results obtained.

The fact that ornithine and arginine are precursors for the polyamines synthesis (Figure 1) led us hypothesize that, polyamines may regulate the levels of ornithine and/or arginine by controlling or regulating the activity of *Xc*AOTC.

To test this hypothesis I assayed the activity of *Xc*AOTC in the presence of several concentrations of putrescine and spermidine (Table 2). The results showed that putrescine did not have an effect over *Xc*AOTC activity, while spermidine behave as an activator of the enzyme reducing the K_m for AOrn and therefore increasing its affinity for this substrate.

Arginine biosynthetic pathway in *Xanthomonads*

The discovery of the AOTC activity of *Xc*ArgF' implied that the following enzyme in the arginine biosynthetic pathway (AOD, Figure 2) should be able to deacetylate acetylcitrulline efficiently in order to continue the flow of metabolites that will produce arginine.

To corroborate this hypothesis I tested the deacetylase activity of *Xc*ArgE using acetylcitrulline and acetylornithine as substrates. The results obtained (Figure 26) showed that the rate at which acetylcitrulline was deacetylated by *Xc*ArgE was comparable to that of acetylornithine deacetylation by the same enzyme. These results probed that acetylcitrulline continues the arginine biosynthetic pathway by means of ArgE and let us propose a novel pathway for arginine biosynthesis occurring in the *Xanthomonads* family of microorganisms and those organisms that possesses an AOTCase.

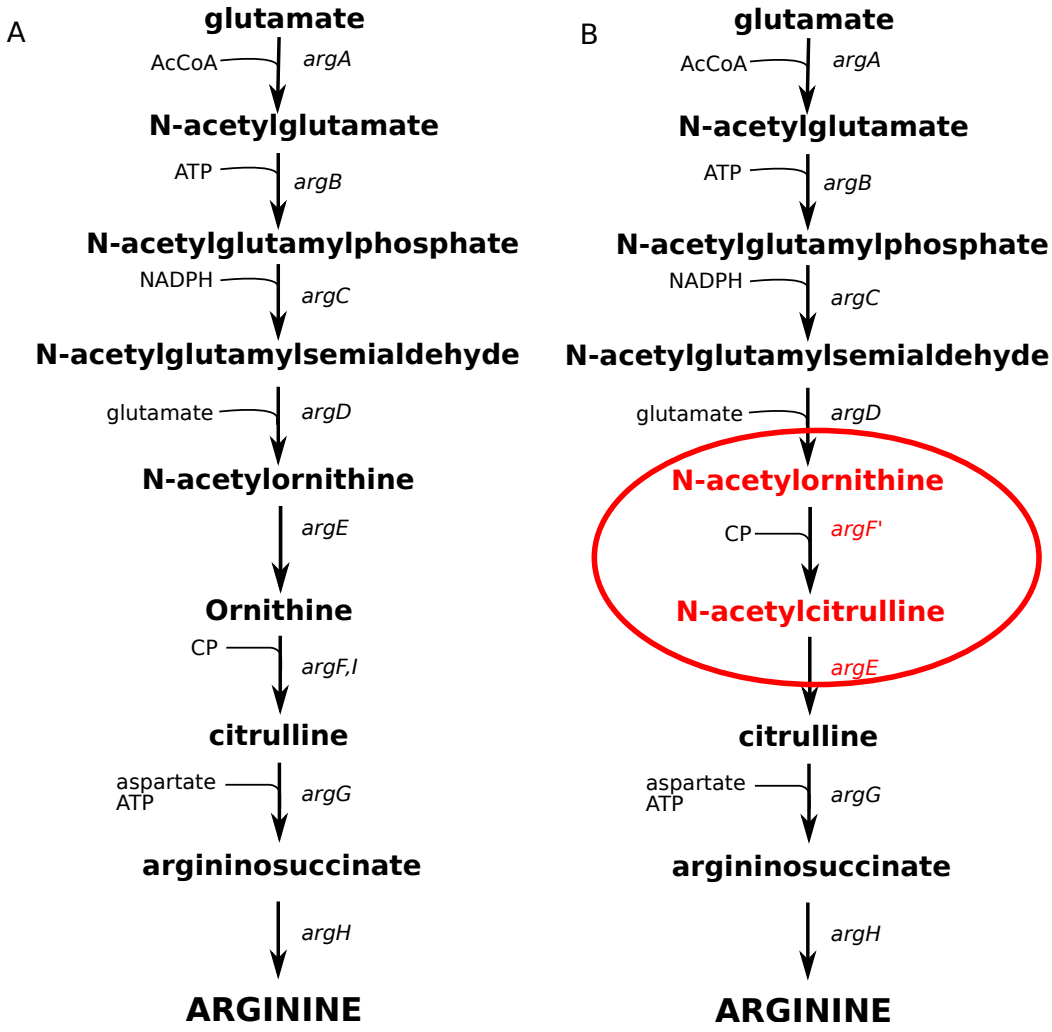


Figure 48: Brief schematic representation showing the proposed arginine biosynthetic pathway for those organisms that contain an AOTC (Panel B) instead of OTC (Panel A). The enzymes studied in this work that differ from the canonical arginine biosynthetic pathway are highlighted in red.

The fact that *Xc*ArgE is able to deacetylate acetylcitrulline as well as acetylornithine implies that a possible regulatory mechanism may occur at the level of ArgE in the arginine biosynthetic pathway of organisms containing an AOTC instead of OTC, which could control the flow of metabolites towards the polyamines or arginine biosynthetic pathways.

Further studies are necessary to corroborate possible regulatory mechanisms of metabolic flow between the arginine and polyamines biosynthetic pathways. To my knowledge such studies are not being performed and could be one of the future perspectives that can be extracted from the present Thesis work.

As part of the research performed in our lab the gene annotated as *argB* in *X. campestris* has been identified as a bifunctional enzyme that is able to synthesize NAG from AcCoA and glutamate and catalyze the ATP-dependent phosphorylation of NAG[82]. Qu et al. have characterized both functions of *Xc*NAGS-K[82] and have concluded that this new bifunctional enzyme activity is regulated by the final product of the pathway, arginine, although the NAGS activity is 25 times more

sensitive to arginine than the NAGK activity.

This recent discovery allows us to re-write the initial part of the proposed arginine biosynthetic pathway for Xanthomonads (Figure 48) to accommodate this finding (Figure 49).

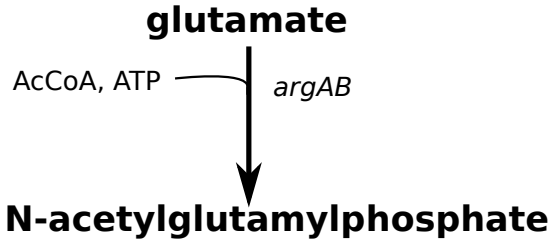


Figure 49: Modification of the initial steps of the arginine biosynthetic pathway in *X. campestris* and organisms that contains AOTC.

Qu et al.[82], on the basis of phylogenetic trees constructed with the sequences from 31 organisms including both bacteria and eukaryotes, speculate that the *argAB* fusion gene arose early in the evolution and underwent gene duplication evolving then to the actual mammalian NAGS and fungal NAGK genes.

Interestingly, the evolutionary tree of the transcarbamylases (Figure 5) generated by Naumoff et al.[46] also locate the organisms containing the *argF"* gene closer to the root than those that contains the canonical *argF* in the evolution of the members of this family of proteins.

The fact that organisms containing the *argF"* gene (which codifies for the enzyme characterized in this Thesis, AOTC) also contains the *argAB* fusion gene suggest that this form of arginine biosynthetic pathway could have been present in the early stages of the evolution.

The discovery of the AOTC activity of *Xc*ArgF" expanded our view of the possible substrates of some of the transcarbamylases of unknown activity identified by Naumoff et al.[46] and by Dr. Morizono.

Implications of this new arginine biosynthetic pathway found in *Xanthomonads*

As stated in the Introduction, ornithine is an intermediate in the synthesis of the following metabolites:

a) Arginine
b) Proline
c) Putrescine
d) Siderophores
e) Peptidoglycan

f) Antibiotics
g) Membrane lipids

In the case of Xanthomonads, AOTC ensures that acetylornithine will continue the arginine biosynthetic pathway; therefore factors altering its activity will affect the availability of ornithine for other pathways.

Metals are an important component of the ecosystem where Xanthomonads can be found. Variations in the concentrations of certain metals could be detected by the cells, reducing the activity of AOTC, while incrementing the activity of ArgE, therefore increasing the concentration of free ornithine inside the cells and activating the synthesis of siderophores. Siderophores are compounds of small molecular mass (<1000 Da) that posses a high affinity for Fe^{2+} and other metals[94], and are usually synthesized in situations were there is a deficiency of iron and is required for cell survival. It has been shown that affecting the iron uptake through the siderophore FeoB, affects the growth (*in planta*) and virulence of *Xanthomonas oryzae pv oryzae*[95] in iron limiting situations therefore it would be interesting to investigate the possible control of the synthesis of siderophores exerted by the regulation of the activities of both AOTC and ArgE in *X. campestris pv campestris* to elucidate a possible mechanism for infections control.

The importance that the regulation of AOTC activity, in response to different environmental situations like heavy rain, temperature and soil pH, has on any of the above mentioned pathways is an interesting project that could be studied if a *X. campestris* ArgF"- and/or ArgE- deficient strains can be obtained.

Discovery and characterization of the enzymatic activity of the protein codified by the gene *argF'* in *Bacteroides fragilis*

The transcarbamylase-like protein codified by the gene *argF"* in *B. fragilis* presented the next challenge in this project. As indicated in the introduction to Chapter 4: "Demonstration that a structurally obvious transcarbamylase from *Bacteroides fragilis*, of previously unknown activity catalyzes the transcarbamylation of succinylornithine. Functional characterization of this transcarbamylase", the crystal structure of this protein has been known since 2002[47] unfortunately, its enzymatic activity remained a mystery as well as the arginine biosynthetic pathway in this organism.

Although the *Bf*ArgF" structure is very similar to that of anabolic OTC and ATCases (Figure 7) there are several differences in some of the regions of the active site, especially at the ornithine binding site, that set the *Bf*ArgF" enzyme apart from previously characterized transcarbamylases. Basically the pocket in which ornithine should bind is bigger in *Bf*ArgF" than in other OTCases (Figure 27 and Figure 50) making impossible for this enzyme to catalyze the transcarbamylation of ornithine or even acetylornithine to produce citrulline or acetylcitrulline (Figure 15).

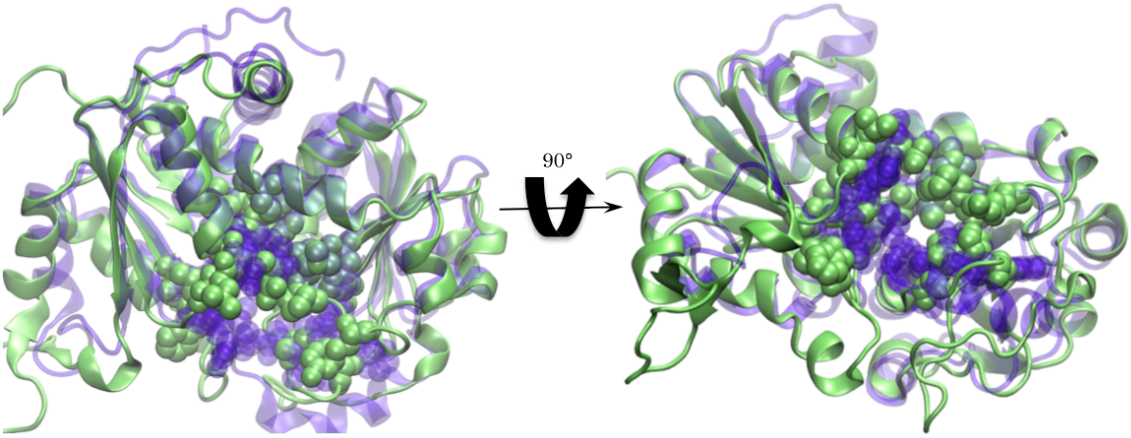


Figure 50: Representation of the binding pockets of human OTCase and *Bf*SOTCase for ornithine and succinylornithine respectively. Human OTCase structure is shown in purple (transparent) and the residues involved in the binding of the substrates are shown as VDW representations. BfSOTCase is shown in green and the residues involved in the binding of the substrates are shown as VDW representations. This figure shows how the substrates-binding pocket in human OTCase (purple transparent) is smaller and less open compared to the substrates-binding pocket in *Bf*SOTCase (green opaque). This figure has been generated using the VMD 1.9 software.

Succinylornithine is an intermediate in the arginine and proline metabolism of several organisms as well as in other biosynthetic and catabolic pathways[96, 97] and its structure is very similar to that of the compounds used as substrates for OTC and AOTC (Figure 28), therefore one of the possible substrates for *Bf*ArgF' is succinylornithine.

To test if *Bf*ArgF' indeed catalyzed the transcarbamylation of succinylornithine to produce succinylcitrulline I chemically synthesized this compound from ornithine and succinic anhydride.

In the first attempt to assay the possible succinylornithine transcarbamylase activity of *Bf*ArgF' we used the lysate of cells over expressing the protein and the results obtained showed a significant amount of activity when succinylornithine was used as substrate for the reaction compared to acetylornithine and ornithine (Figure 29). These results led us to hypothesize that we had found the activity of *Bf*ArgF' and is that of a succinylornithine transcabamylase (SOTC) and therefore yet another new member of the transcarbamylases family of enzymes.

To confirm this hypothesis and to discriminate between possible artifacts and real SOTC activity from the over expressed *Bf*ArgF' in the lysates, we purified *Bf*ArgF' and tracked the SOTC activity throughout the purification (Figure 30). The results showed that the SOTC activity was detected in those fractions in which *Bf*ArgF' was present. From these experiments we confirmed the initial hypothesis and concluded that the transcarbamylase activity of *Bf*ArgF' was that of a succinylornithine transcarbamylase, a new member of the transcarbamylases family of enzymes.

Further prove of the SOTC nature of *Bf*ArgF' was obtained from the NMR and LC-MS analyses of the product of the reaction (Figure 31). These analyses confirmed that the product of the reaction was indeed N-succinylcitrulline.

Once that we proved that *Bf*ArgF' is indeed a SOTC we proceeded to characterize its activity by determining the kinetics parameters K_m and V_{max} for both substrates: SOrn and CP. A summary of the kinetic parameters determined for

both substrates of *Bf*SOTC can be found on Table 4. Although the affinity of *Bf*SOTC for CP is on the range of other transcarbamylases[41, 69], its affinity for SOrn is the highest among the studied transcarbamylases. Such a high affinity could be explained by an error on the estimation of the K_m value for SOrn due to the limited amount of concentrations studied.

The SOrn saturation curves showed that *Bf*SOTC is slightly inhibited by high concentrations of this substrate (Figure 33), while CP saturation curves followed a Michaelis-Menten kinetic (Figure 34). Structural insights on the conformational changes occurring in *Bf*SOTC upon binding of the substrates provided an explanation for the substrate inhibition detected[85]. From the studies performed on the crystallographic structure of *Bf*SOTC complexed with its substrates it was possible to quantify the domain closure of the enzyme upon binding of the substrates. In the case of *Bf*SOTC when CP and succinylnorvaline were bound, the domain movement was of ~5.6°[85] a value comparable to that of *E. coli* and human OTCases is of 4-12°[52, 91], while in *E. coli* ATC is 8.4-12.3°[98] and in *Xc*AOTC is only 2.2°[63]. As stated before, the catalytic mechanism of *Xc*AOTC is random which explains the lack of substrate inhibition, while in OTC and ATCases is ordered[52, 69, 87, 88] as in *Bf*SOTC, differences in the interactions of the substrates to each enzyme can explain the differences in the catalytic properties observed between them.

The reaction catalyzed by *Bf*SOTC is very specific for succinylornithine as has been demonstrated in the results presented in this Thesis (Figures 15 and 29), which led us to think that succinylation would be required for inhibitors recognition as well. PALO and its acetylated derivative PALAO are two potent inhibitors of OTC and AOTC respectively, therefore it was expected that the succinyl derivative of PALO αN-succinyl-δN-phosphonoacetyl-L-ornithine (PALSO), would be also a strong inhibitor of *Bf*SOTC.

To test this hypothesis I synthesized PALSO chemically from PALO in the presence of succinic anhydride. *Bf*SOTC activity was assayed in the presence of several concentrations of PALSO (Figure 35) and estimated IC_{50} was 0.45 ± 0.05 μM. Although the PALSO preparation had a very high degree of purity (Figure 35), due to the lack of purified PALSO we could only estimate the IC_{50} and not calculate the K_i or determine the mechanism of inhibition of this compound over *Bf*SOTC activity.

Arginine biosynthetic pathway in *Bacteroides fragilis*

The fact that *Bf*ArgF' was unable to complement OTC deficient mutants of E. coli, while restoring the capacity to synthesize arginine in the *B. fragilis* arginine auxotrophs JDP3a.3 or JDP3a.4[85], along with the discovery of its SOTC activity, implies that the pathway by which *B. fragilis* and, therefore the *Bacteroidetes* family of microorganisms synthesize arginine is different from any organism described until now. We hypothesized that, in *Bacteroidetes*, arginine is synthesized by a pathway that use succinylated rather than acetylated intermediates, and to probe it I investigated the initial steps of the pathway.

There is no gene similar to *argA* in the whole genome of *B. fragilis* although looking at the genomic organization of the genes involved in the arginine biosynthesis in this organism it is possible that the gene denominated *argX?* (Figure 37) could encode the enzyme responsible for the synthesis of succinylglutamate if the

hypothesis of a succinylated arginine biosynthetic pathway is true.

I cloned and over expressed the *argX* gene from *B. fragilis*, unfortunately its activity could not be assayed due to the high rate of spontaneous succinylation of glutamate in the presence of Suc-CoA (Figures 38-40).

Therefore I focused my efforts on the next enzyme of the pathway, ArgB (Figure 2). In the arginine biosynthetic pathway known until now, N-acetylglutamate kinase (NAGK or ArgB) catalyzes the ATP-dependent phosphorylation of NAG, in the second step of the pathway (Figure 2)[9]. In the case of *B. fragilis*, hypothetically, ArgB would catalyze the ATP-dependent phosphorylation of N-succinylglutamate (NSG).

Once again, the substrate of the reaction that I was trying to assay was not commercially available therefore I synthesized it chemically from glutamate and succinic anhydride. The degree of conversion was very high (99.95%, based on LC-MS quantification of the free glutamate present in the sample after the chemical reaction was stopped) and the chemically synthesized NSG was used without further purification to test the NSGK activity of *Bf*ArgB.

The initial attempts to assay the activity of *Bf*ArgB were performed using lysates of cell over-expressing the protein (Figure 42) by a colorimetric method and showed an significant increase in the color produced in the sample that contained chemically synthesized NSG with respect to the other substrates assayed (NAG, glutamate, succinylaspartate and aspartate). These results indicated that, most likely, the activity of *Bf*ArgB is that of an N-succinylglutamate kinase.

To probe that NSGK was indeed the activity of the enzyme codified by the gene *argB* in *B. fragilis*, I purified *Bf*ArgB with a degree of purity >95% and assayed it in the presence of the same substrates as the lysates. The results showed that NSGK was the preferred substrate since the partial activity detected with NAG and succinylaspartate was not distinguishable from the heat-inactivated enzyme controls (Figure 43).

Using the chemically synthesized NSG I proceeded to determine the kinetic parameters that characterized *Bf*NSGK using NSG and ATP saturation curves. The K_m and V_{max} determined for each of the substrates of *Bf*NSGK are summarized on Table 5.

The affinity of the enzyme for the two substrates apparently is lower compared to other NAGKs from other organisms that uses use acetylated intermediates in their arginine biosynthetic pathway[83, 84], nevertheless chemically synthesized and purified NSG should be obtained in order to properly determine the kinetic parameters of *Bf*NSGK. Proximity of the substrates would help to compensate for the low affinity of *Bf*ArgB for them, therefore it is possible that ArgB and ArgX from *B. fragilis* interact with each other in such a way that ArgX becomes active and the substrates are easily available to ArgB to continue the arginine biosynthesis pathway as is the case of the *S. cerevisiae* and *N. crassa* enzymes[17].

The discovery of the SOTC activity of *Bf*ArgF' implied the existence of a succinylated pathway for arginine biosynthesis in this organism. Complementation studies performed by Dr. Malamy provided further evidences that the arginine biosynthetic pathway in *B. fragilis* was different from that of *E. coli*. Although the enzyme responsible for the synthesis of NSG in *B. fragilis* remains elusive, we have demonstrated that the second step of the arginine biosynthetic pathway in this organism, catalyzed by the product of the gene *argB* requires NSG, confirming the

hypothesized succinyl derivatives intermediated pathway for arginine biosynthesis in *B. fragilis* and other organisms that contains SOTC.

The fact that there is no OAT-like gene in the whole genome of *B. fragilis* and the fact that arginine has no regulatory effects on the activity of $_{Bf}$ArgB led us to hypothesize that the arginine biosynthetic pathway in *B. fragilis* and organisms with SOTC is linear instead of cyclic. A schematic representation of the proposed arginine biosynthetic pathway occurring in Bacteroidetes is shown in Figure 51.

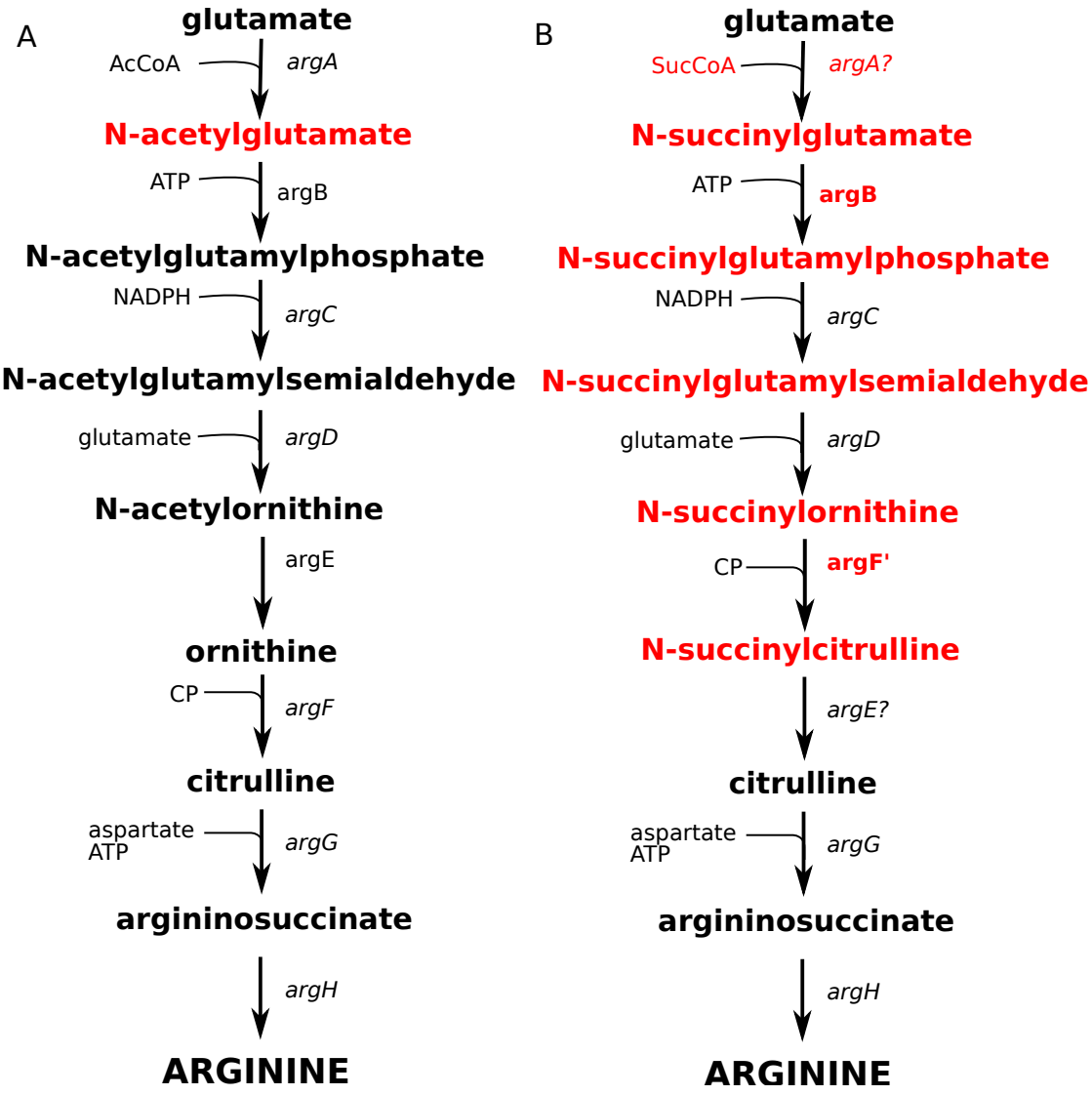


Figure 51: Schematic representation of the proposed arginine biosynthetic pathway occurring in Bacteroidetes and other organisms with SOTC (Panel B) compared to that occurring in organisms with OTC (Panel A). For clarification, differences between the two arginine biosynthetic pathways are shown highlighted in red.

Structural insights onto substrate recognition of *Xc*AOTC and *Bf*SOTC

Comparison of the crystal structures from *Xc*AOTC and *Bf*SOTC highlighted the differences between these two enzymes[86]. Although the tridimensional structures are very similar there are several differences in important regions of the proteins (Figure 52):

- The conformation of the 240's loop is 4 amino acids longer in *Xc*AOTC vs. *Bf*SOTC.
- In the 80's loop of *Xc*AOTC the NE atom of Trp77 (Trp72 in *B. fragilis*) is displaced 12° relative to the substrate. The loop movement is important for CP binding.
- Sequence differences around helix H1 significantly shift its position.
- The major difference between these two structures is an 11 amino acids insertion between H7 and B8, which makes the β-sheet core in the second substrate domain in *X. campestris* more similar to the *E. coli* OTC.

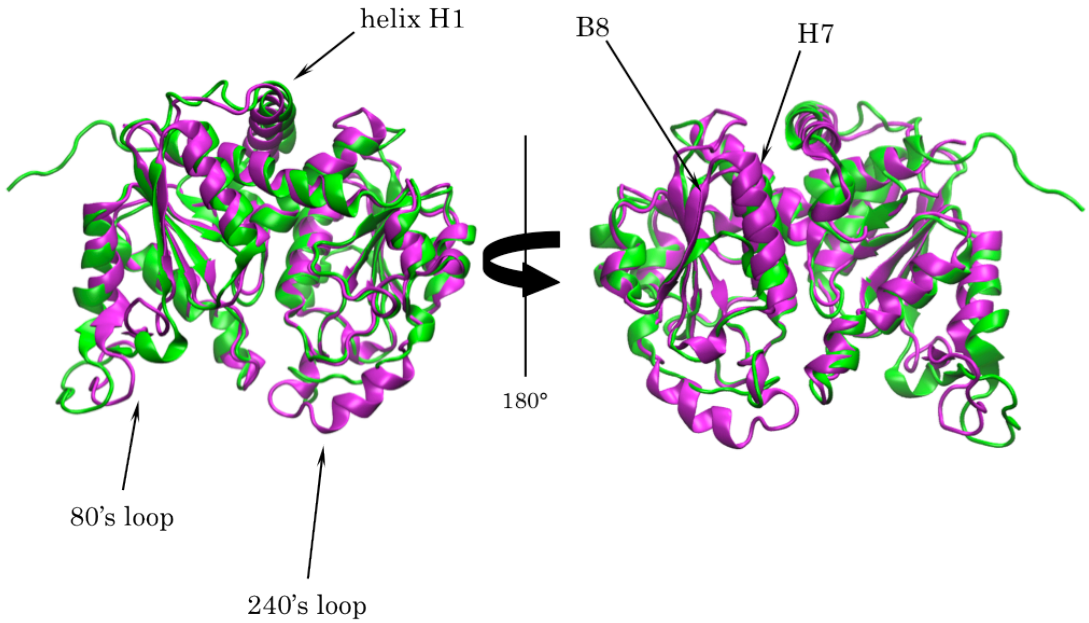


Figure 52: Comparison of the crystallographic structures of *Xc*AOTC and *Bf*SOTC. *Xc*AOTCase strucutre is shown in purple and *Bf*SOTCase structure in green. Differences mentioned in the text are indicated by arrows.

From the crystallographic studies of these two enzymes (*Xc*AOTC and *Bf*SOTC) it was determined that the residue Glu92 (Pro90 in *B. fragilis*) was the responsible for the specificity of the *X. campestris* ArgF' for AOrn[86]. The fact that proline is a smaller residue compared to glutamate was hypothesized to help the fitting of bulkier acyl groups. Around the same time, the activity of *Bf*SOTC was discovered confirming the hypothesis of a substrate bigger than ornithine or acetylornithine. Crystallographic studies of *Bf*SOTC complexed with its substrates and substrates analog highlighted the importance of the residue Pro90 (*B. fragilis* numbering) on the specificity of the enzyme[85].

The importance of this residue (Pro90) in substrate recognition has been studied using the following mutants[99]:

*Xc*AOTC mutants: E92S, E92A, E92V and E92P

*Bf*SOTC mutants: T242L, P90E/T242L

The fact that all of the *Xc*AOTC mutants lost, almost completely, the AOTC activity while catalyzing the transcarbamylation of SOrn to SCit and the fact that the P90E/T242L mutant of *Bf*SOTC was able to catalyze the transcarbamylation of AOrn almost 10 times more efficiently than SOrn[99] confirmed that the residue Pro90 (in *B. fragilis* numeration) or Glu92 (in *X. campestris*) is the responsible for determining the substrate specificity in these two enzymes (*Xc*AOTC and *Bf*SOTC).

As a consequence of the results discussed above these new transcarbamylases activities can be assigned to new genes, based on the presence or absence of glutamate and allowed the proposal of the following schematic method for classification:

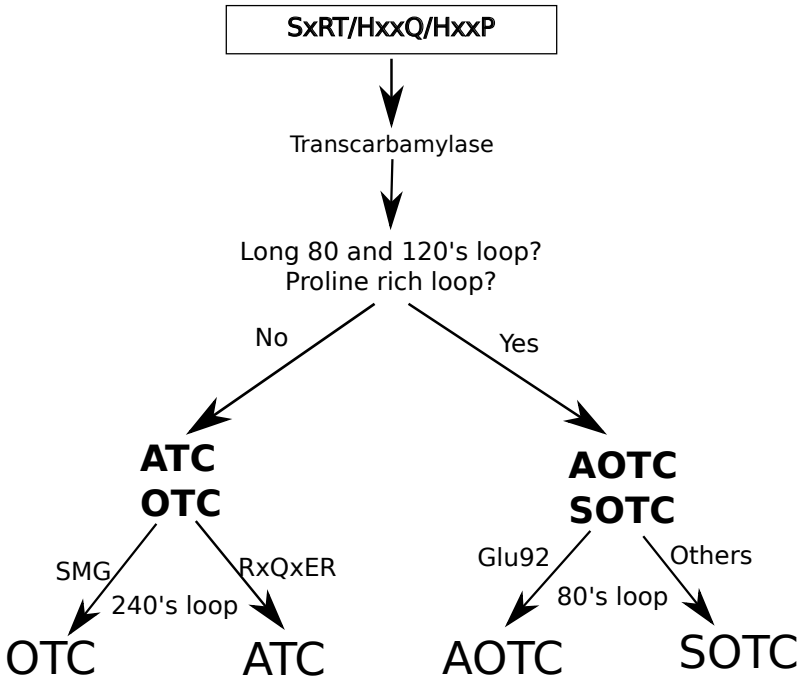


Figure 53: Schematic classification method proposed for ornithine, aspartate, acetylornithine and succinylornithine transcarbamylases based on structure-derived amino acid signature motifs.

Using this classification method Shi et al.[99] proposed the re-annotation of several misannotated transcarbamylases in the genomes of organisms that included several species of the genus *Xanthomonas*, *Xylella* and *Flavobacteria* among others some of which had been confirmed enzymatically.

Can NSGKs also be correctly annotated based on the genomic sequence?

Despite the fact that *E. coli* ArgB and *B. fragilis* ArgB are very similar (47% similarity, 27% identity) there are striking differences between the two proteins as shown in Figure 54.

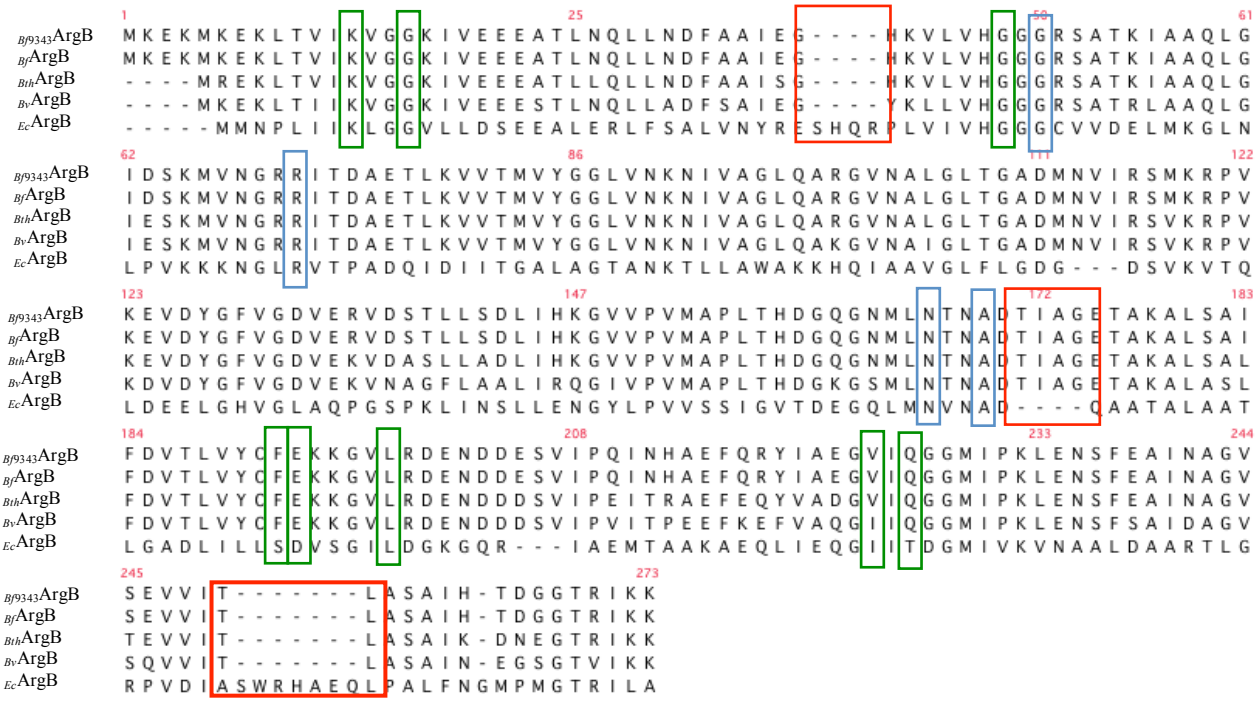


Figure 54: Comparison of the amino acids sequence of *E. coli* ArgB with several members of the Bacteroidetes family. Sequences were obtained from the NCBI Protein database. Sequences alignment was performed using ClustalW. Green boxes indicate residues responsible for ATP binding. Blue boxes indicate residues responsible for NAG binding. Red boxes indicate sequences gaps between *E. coli* sequence and the other proteins used for the analysis.

The 4 residues gap before G45 (*E. coli* numbering) and the 3 residues insert after A161 could explain how a bigger pocket is present in this enzyme to accommodate the larger NSG (compared to NAG), but crystallographic studies are necessary to corroborate this hypothesis.

Some of the residues identified by Ramón-Maiques et al.[100] as responsible for ATP binding (based on the cristal structure of *E. coli* ArgB liganded to AMPPNP) are different in the Bacteroidetes ArgB (S180, D181 and T211, *E. coli* numbering), the relevance of these changes still needs to be determined using site directed mutagenesis and enzymatic studies.

Although, without the proper crystallographic data and the proper enzymatic studies it would be hard to distinguish between NSGK and NAGK based on the amino acids sequence, we can speculate that the above mentioned gaps and insertions combined with the amino acids changes observed in the Bacteroidetes ArgBs could be an useful tool for discrimination between these two types of enzymes.

Conclusions

The conclusions of the present thesis work are the following.

1. The AOTCase from *Xanthomonas campestris pv campestris* characterized in this thesis, catalyzes the formation of a new intermediate in a novel arginine biosynthetic pathway found in some of the bacteria that do not possess the canonical *argF* gene.

2. *Bacteroides fragilis argF'* gene encodes for another new member of the transcarbamylases proteins; N-Succinyl-L-Ornithine transcarbamylase (SOTCase).

3. *Bacteroides fragilis argB* gene encodes for a new enzyme in the arginine biosynthetic pathway is these microorganisms, N-Succinyl-L-Glutamate kinase.

4. The arginine biosynthetic pathway in the Bacteroidetes family of microorganisms uses succinylated instead of acetylated intermediates in the first part the pathway.

5. The identification and biochemical characterization of the transcarbamylases found in *X. campestris* and *B. fragilis* enables the bioinformatic annotation of these transcarbamylase-like genes in other organisms.

Materials and Methods

Materials

Chemicals:

All chemicals utilized in this work were purchased from Sigma-Aldrich or Fischer Scientific unless otherwise specified.

Strains and plasmids:

Tables 6 and 7 contain the cell strains and the plasmids used in the present work.

Table 6: Cell strains

Strain	Characteristics	Purpose	Reference
BL21(DE3)	F– ompT gal dcm lon hsdSB(rB-mB-) λ dST [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	general purpose expression host	Studier et al. 1986
DH5a	F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG lasmlacZΔacZAΔalacZYA-argF)U169, hsdR17(rK- mK+), λ–	non-expression host; general purpose cloning; plasmid propagation.	Messelson et al. 1968
TOP10	F- mcrA Δmmrr-hsdRMS-mcrBC) φrr-hsdRMS- Δrr-hsd nupG recA1 araD139 ΔnupG recA1 ar galE15 galK16 rpsL(StrR) endA1 λ-	host for general purpose cloning and plasmid propagation. expression host used in combination with vectors containing BAD promoter.	Invitrogen
Rossetta2	F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE2 (CamR)	expression host; allows expression of genes encoding tRNAs for rare arginine codons AGA, AGG, and CGA, glycine codon GGA, isoleucine codon AUA, leucine codon CUA, and proline codon CCC	EMD Biosciences
HMS174(DE3)	F- recA1 hsdR(rK12- mK12+) (DE3) (Rif R)	recA- expression host	Novagen
B. fragilis TM4000	Wild type strain of B. fragilis		Dr. Malamy (Tufts University, Boston, MA)
B. fragilis TM4000 4.9	B. fragilis ArgE- mutant		Dr. Malamy (Tufts University, Boston, MA)

Table 7: Plasmids

Plasmid	Purpose	Bacterial resistance	Reference
pCRBluntII TOPO	Cloning	Kanamycin	Invitrogen
pCR4Blunt-TOPO	Cloning	Ampicillin/Kanamycin	Invitrogen
pET15b	Cloning/expression	Ampicillin	Novagen
pET21b(+)	Expression	Ampicillin	Novagen
pET28a(+)	Cloning/Expression	Kanamycin	Novagen
pJCL7	Derived from pET21b(+) containing BfArgB	Ampicillin	This work
pET21b-BfArgX	Derived from pET21b(+) containing putative BfArgA	Ampicillin	This work
pTrc99a-BfArgF-His	Derived from pTrc99a containing BfArgF'	Ampicillin	Dr. Malamy

Media and growth conditions

Liquid Media

The rich liquid medium used for the growth of all the strains used in this work was the Luria-Bertani (LB) medium. The following is the composition of this media:

Bactotriptone	10 g
Yeast extract	5 g
NaCl	5 g
H_2O	up to 1 liter

Media were sterilized by vapor pressure in the autoclave at 121°C during 20 minutes.

Antibiotics were added to the cooled, sterilized media at the concentrations specified in Table 8.

Table 8: Antibiotics

Antibiotic	Stock concentration	Working concentration	Solvent
Ampicillin (Amp)	60 mg/ml	60 µg/ml	H2O
Kanamycin (Kan)	10 mg/ml	10 µg/ml	H2O
Chloramphenicol (Cam)	35 mg/ml	35 µg/ml	Ethanol

Solid Media

For the preparation of the solid media, liquid medium (LB) was supplemented with bacto-agar to a concentration of 1.5% (w/v). In this case the antibiotics were added after sterilization and before the media solidified (approximately at a temperature of 50°C).

Plasmid isolation

For plasmids isolation strains containing the plasmid to isolate were grown in 3 ml or 100 ml of LB supplemented with the required antibiotic, for mini-prep or midi-prep respectively, overnight (o/n) at 37°C under aerobic conditions. The cultures were shaken at 250rpm in an Innova 4000 (New Brunswick Scientific) orbital shaker.

For strain conservation 1 ml of the o/n culture was mixed with 1 ml of sterile 50% glycerol and kept at -80°C.

Protein expression

For protein expression, strains were grown o/n in 50 ml of liquid LB supplemented with the required antibiotic, at 37°C and 250 rpm. The following day, 1 L of LB supplemented with the required antibiotic was inoculated with 20-40 ml of the o/n culture and cells were grown at room temperature (r.t.) until OD_{600} reached 0.4-0.6 absorbance units. When the culture reached the desired OD_{600}, recombinant protein expression was achieved by adding IPTG to the medium to a final concentration of 0.1-0.2 mM. The cells were grown o/n under expression conditions.

Buffers

Transcarbamylases purification buffers

Table 9 indicates the composition of the buffers utilized for the purification of the transcarbamylases utilized in this work.

Table 9: Transcarbamylase purification buffers

Column	Binding buffer	Elution buffer
Nickel column	50mM NaH2PO4 300mM NaCl 10% Glycerol 10mM b-ME pH 7.4	50mM NaH2PO4 300mM NaCl 10% Glycerol 10mM b-ME 250mM Imidazole pH 7.4
DEAE column	20mM Tris-HCl 50mM NaCl 1mM EDTA 5mM b-ME pH 8.0	20mM Tris-HCl 500mM NaCl 1mM EDTA 5mM b-ME pH 8.0

Transcarbamylases storage buffer

Bacteroides fragilis and *Xanthomonas campestris* pv *campestris* transcarbamylases were stored in DEAE column binding buffer.

NAGS purification buffers

Table 10 indicates the composition of the buffers utilized for the purification of the NAGS-like proteins utilized in this work.

Table 10: NAGS purification buffers

Column	Binding buffer	Elution buffer
Nickel column	50mM KPO4 300mM KCl 10mM Imidazole 20% glycerol 0.006% Triton X-100 pH 7.5	50mM KPO4 300mM KCl 500mM Imidazole 20% glycerol 0.006% Triton X-100 pH 7.5
DEAE column	50 mM Tris 50mM NaCl 20% Glycerol 10mM b-ME 0.006% Triton X-100 pH 8.0	50 mM Tris 1M NaCl 20% Glycerol 10mM b-ME 0.006% Triton X-100 pH 8.0

NAGS storage buffer

NAGS-like proteins were stored at 4°C in the following buffer:

50mM KPO_4, pH 7.5
50mM KCl
10mM β-ME
0.1% Triton X-100
20% Glycerol

Tris-Acetate-EDTA (TAE) buffer

50X TAE stock solution was prepared as follows:

242 g Tris Base (MW=121.1)
57.1 mL Glacial Acetic Acid
100 mL 0.5 M EDTA
Add H2O to a final volume of 1L
Store at room temperature.

Note: Final (1x) working concentration: 0.04 M Tris-Acetate, 0.001 M EDTA

6x DNA loading buffer

0.25% Bromophenol Blue
0430% Glycerol

Methods

Primer design

Primers were designed using the software package Amplify 3© (Bill Engels, University of Wisconsin. http://engels.genetics.wisc.edu/amplify)

Table 11 shows the primers designed for the completion of the present thesis work.

Table 11: Primers

Primer	Sequence (5'-3')
BfargBNdeIFW	TTTCATATGAAAGAAAAGATGAAGGAAAAACTGAC
BfragBXhoI	TTTCTCGAGTTTTTTTATCCGAGTTCCCCCATCT
BfargXXhoIFW	ATTCTCGAGGTTATTTCTTTCTTGTTCTTTTTTTATGTTA
BfargXNdeIRV	TCACATATGGACACTCAACAAATAGATGTTATGGTAG

Cloning of *Bf*ArgB

The sequence flanking argB gene was identified from genomic sequences of *B. fragilis* NCTC9343 using Artemis® software, and used to design the primers BfargABNdeIFW and BfragABXhoI that add the restriction sites of NdeI and XhoI on the 5' and 3' ends of the PCR amplification product, respectively. A crude lysate of *B. fragilis* TM4000 4.9 (argE$^-$) was used as template for PCR, and argB was amplified using Phusion DNA polymerase following manufacturer instructions. The cycling conditions were as follows: 30 s at 98°C for polymerase activation and 40 cycles for the melting (10 s, 98°C), annealing (15 s, 55°C) and extension (15 s, 72°C), followed by an extension step (5 min, 72°C) yielding an unique fragment of the desired size (798bp). The 798bp-long PCR amplification product was cloned into pCRBluntII vector (Invitrogen), subcloned into the NdeI and XhoI sites of pET21b (Novagen), and sequence verified to produce the pJCL7 plasmid.

Cloning of the putative *Bf*ArgA

Cloning of the putative *Bf*ArgA (ORF# BF0842 in *B. fragilis* NCTC9343 strain genomic sequence), was performed in a similar way as the cloning of *Bf*ArgB but using the primers: BfargXXhoIFW and BfargXNdeIRV for amplification. Cloning of the *Bf*ArgA yielded the expression vector pET21b-*Bf*ArgX.

Transformation

One aliquot of 50µl of chemically competent cells (Table 8) were transformed with 1µl of the desired plasmid (Table 9) at ~10ng/µl, followed by a 10 minutes incubation on ice. After the incubation on ice cells were heat shocked at 42°C for 30 seconds and immediately were placed on ice. 200µl of S.O.C. medium were added to

the transformed cells and then incubated at 37°C and 250 r.p.m. for 1 hour. After incubation at 37°C 10, 25 and 50µl of the transformed cells were plated on LB-agar containing the required selection antibiotics (Table 8).

Plasmid DNA purification

Plasmids were purified using the QIAprep Spin Miniprep Kit or the QIAGEN Plasmid Midi Kit (QIAGEN) following the manufacturer protocol.

DNA digestion with restriction enzymes

All the restriction enzyme reactions were carried out using enzymes purchased from New England Biolabs (NEB) following the manufacturer protocols using 0.5 to 1µg of plasmid DNA. When possible, double digestions reactions were performed.

DNA electrophoresis

Following PCR amplification and enzymatic restriction reactions, DNA samples were separated based on size using horizontal electrophoresis in TAE buffer on agarose gels. Agarose gels were prepared using 1x TAE buffer and depending on the size of the fragment to identify the composition of the agarose gel varied from 0.9 to 2% w/v. 10µl of sample were mixed with 2µl of 6x loading buffer prior to the load of the sample on the gel.

Once loaded on the gel samples were separated by electrophoresis at 90 volts for 45 minutes, or 100 volts for 60 minutes, depending on the size and composition of the gel and the number of samples

1Kb(+) (Invitrogen) size ladder was used to identify the size of the separated fragments.

Sequencing

Purified plasmids and PCR product were sequenced using the services of ACTG Inc. (Wheeling, Illinois) and following the company guidelines for sample preparation and concentration.

Sequencing results were analyzed using Sequencher® software package (Genes Codes Corporation).

Protein purification

Cells were harvested by centrifugation, suspended in 40 ml of sonication buffer (Ni-column buffer A) and lysed by sonication using a Bransonic Sonifier at approximately 50% maximal output for 10 minutes on ice. The lysates were cleared by centrifugation at 14,000-x g for 25 minutes. The supernatant fluid was filtered

through a 0.22 µm filter and then loaded at a flow rate of 1 ml/min onto a 5 ml His-Trap Ni-affinity column (Amersham Biosciences) equilibrated with Ni-column buffer A. The eluted volume was monitored for the presence of proteins by UV-spectrophotometry at 280 nm.

Weakly bound or unbound proteins were removed by washing the column with the same buffer containing 50 mM imidazole (Ni-column buffer B at 20% proportion) until the absorbance at 280 nm returned to baseline. The imidazole concentration was then increased to 250 mM to elute the protein of interest. Protein-containing fractions were pooled; dialyzed into DEAE-column buffer A; and then loaded onto a DEAE His-Trap column equilibrated with the same buffer running at 3.0 ml/min. A linear gradient of 20 column volumes with DEAE-column buffer B as second buffer was used to elute the protein. The eluted protein was dialyzed and stored in DEAE-column buffer A.

Protein concentration

The purified protein was concentrated using a Millipore 30,000 or 100,000-molecular-weight-cutoff spin filter to > 10 mg/ml and stored at 4°C.

Protein quantification

For all procedures, protein concentration was determined by the Bradford method[101] using the Bio-Rad protein assay dye reagent with bovine serum albumin (BSA) as standard.

Protein electrophoresis (SDS-PAGE)

SDS-PAGE was used to separate the proteins of each preparation according to their molecular weight. 0.7-20ng of total protein were mixed 1:1 with Laemmli sample buffer (Bio-Rad) prepared according to the manufacturer's instructions. Samples were then incubated in boiling water for about 10 minutes. After that, samples were loaded onto the appropriate Ready Gel® or Criterion® precast gels (Bio-Rad) depending on the number of samples. Electrophoresis was performed at 120 volts using a PowerPac® power supply (Bio-Rad) for 1 hour and stained with Bio-Safe Coomasie Stain (Bio-Rad) according to manufacturer instructions.

Colorimetric transcarbamylase activity assay

Transcarbamylase activity was measured by using a modification of the colorimetric assay to detect an ureido group[64], except that AORn or SOrn replaced ornithine. All assays were performed in triplicates at 25°C. The reaction was initiated with CP, allowed to proceed for 5 minutes and then quenched by the addition of 1 ml of antipyrine:butadione monoxime color reagent. Fresh color reagent (1 part of 0.8% 2,3-butadione monoxime [wt/vol] dissolved in 5% acetic acid [vol/vol] mixed with 2 parts of 0.5% antipyrine dissolved in 9.05 M H_2SO_4) kept on

ice was used to terminate the reactions. The assay is very sensitive to light; therefore, samples were then wrapped in foil and placed on a shaker, also protected from light, overnight at approximately 100 rpm. To develop the final color, the samples were uncovered and incubated for 20 minutes at 42°C with illumination. The chromophore in the assay produced by commercially synthesized and purified ACit has an absorption coeffient that was 93.4% of that of Cit.

Standard curves of known amounts of ACit were used for quantification of the enzymatic activity.

Initial reaction mixtures to determine the function of the ArgF' protein from *Xanthomonas campestris* contained 4 mM of AOrn and 4.8 mM CP in 1 ml of a 50 mM Tris-HCl, pH 8.3 buffer.

Colorimetric NAGS activity assay

ArgB kinase activity was measured using a colorimetric assay described by Haas and Leisinger [84], to detect acetylglutamate kinase activity but substituting N-acetyl-L-glutamate (NAG) by N-succinyl-L-glutamate (NSG). Succinylglutamyl hydroxamate is formed from NSG-phosphate in the presence of ammonium hydroxamate. The 500 µl incubation volume contained 100 mM MES buffer pH 6.0, 40 mM ATP, 40 mM NSG, 400 mM $NH_2OH.HCl$, 50 mM MgCl2 and 2.5 µg of protein. The reaction mixture was incubated for 10 min at 37°C and terminated by the addition of 500 µl of 5%(w/v) ferric chloride solution ($FeCl_3 \cdot 6H_2O$) containing 8%(w/v) trichloroacetic acid and 0.3 M HCl. The absorbance of the hydroxamate–Fe^{3+} complex was measured at 540 nm.

Saturation curves

Activity determinations for the kinetic studies were performed by measuring the amount of ACit or SCit formed at 25°C in a 1 ml reaction mixture in the presence of 0.2 µg of enzyme as described before. The rate of catalysis of the enzymes was investigated using saturation curves for each of the substrates of the enzyme.

$_{Xc}$AOTCase saturation curves: In the AOrn saturation curves, CP was kept constant at a concentration of 4.8 mM, while the AOrn concentration was varied from 0 to 5 mM, the reaction was started by adding CP to the mixture. K_m and V_{max} were determined from the plots of v vs. [AOrn], and non-linear fitting using the Michaelis-Menten equation for enzyme kinetics. The same procedure was used to obtain the kinetics parameters of the enzyme for CP, in this case AOrn concentration was kept constant at 5 mM and that of CP was varied from 0 to 1 mM, and the reaction was initiated by adding AOrn to the mixture.

$_{Bf}$SOTCase curves were obtained as the $_{Xc}$AOTCase saturation curves but SOrn was the substrate and SCit was used as standard reference.

In the case of $_{Bf}$ArgB the substrates were ATP or NSG and the saturation curves were obtained as descrived above.

Liquid chromatography mass spectrometry (LC-MS)

Samples were treated with equal volumes of 30% Trichloroacetic acid (TCA) and centrifuged for 10 minutes at 14,000-x g to remove the precipitated protein. 10-20µl of the reaction solution were analyzed using an Agilent LC-MS 1100 (Hewlett-Packard) and conditions described previously by Caldovic et al.[102], briefly: compounds are separated by reverse-phase HPLC (Hewlett-Packard 1100). The mobile phase consisted of 92% solvent A (1ml of trifluoroacetic acid in 1L of water) and 8% solvent B (1ml trifluoroacetic acid in 1L of 1:9 water:acetonitrile). The flow rate was 0.6ml/min. Compounds were quantified using the area under the selected ion peak extracted after mass spectrometry analysis (1100 MSD, Hewlett-Packard).

Nuclear magnetic resonance (NMR)

Enzymatically synthesized ACit and SCit were analyzed using NMR as described in this section.

Three milligrams of purified ArgF' protein were transferred into buffer containing 50 mM sodium phosphate and 100 mM NaCl, pH 7.4, using three rounds of centrifugation with a Millipore Microcon YM30 spin filter according to manufacturer's directions. A total of 15.4 µg of the buffer-exchanged protein were reacted with 4.8 mM CP and 5 mM AOrn or SOrn dissolved in the same buffer at pH 8.3. After overnight incubation at 25°C, the enzyme was removed by centrifugation through a Millipore Microcon YM30 filter. Two micrograms of CP, AOrn, SOrn and chemically synthesized ACit or SCit were dissolved in the phosphate buffer used for the assay for use as standards. The enzymatically and chemically synthesized ACit, SCit, SOrn, CP and AOrn were subjected to three rounds of lyophilization and resuspension in D_2O prior to analysis to suppress the residual proton (^1H) water signals.

One-dimensional ^1H NMR was performed using a Bruker NMR spectrometer (DRX-500; Bruker, Billerica, MA) equipped with a tri-nuclei inverse probe with a Z gradient at 500.13 MHz at 25°C ± 0.5°C. Chemical shifts were calibrated with an external reference of 3-trimethylsilyl propanesulfonic acid, sodium salt, set as 0 ppm at the same temperature. NMR line assignments were based on the similarities (within 2 Hz) in the one-dimensional ^1H chemical shifts between the model compounds (chemically synthesized ACit, SCit, SOrn and AOrn) and the enzymatically-synthesized product at the same pH, ionic strength and temperature. Two-dimensional NMR-correlated spectroscopy (COSY) with solvent presaturation was carried out using a Bruker NMR spectrometer (AV-400) fitted with a broadband inverse probe with a Z gradient at 400.13 MHz at 25°C. The relaxation delay was 2.5 s. The two-dimensional matrix for COSY was 2000 by 512, yielding digital-free induction decay resolutions of 2.1 Hz and 8.6 Hz, respectively. Linear prediction was used for processing of the data, giving a matrix size of 2000 by 1000.

NMR analyses were performed by Yui-Fai Lam at the University of Maryland NMR and EPR Analytical Facility.

Ninhydrin reactivity assay

The product of the reaction catalyzed by *X. campestris* ArgF' was analyzed using the ninhydrin reactivity assay.

Ninhydrin preferentially reacts with free amino groups to produce a detectable product with a maximal absorbance at 570 nm. 10 µg of purified *X. campestris* ArgF were incubated for 5 minutes at 25°C with 5 mM AOrn and 4.8 mM CP in a 1-ml volume. 100 µl of the reaction mixture were removed and diluted to 500 µl with water, and the reaction was terminated with 500 µl of 7% sulfosalicylic acid. 900 µl of color reagent were added to the remainder to terminate the reaction and for quantification of the product by the colorimetric assay. The sulfosalicylic acid-treated sample was analyzed using a Beckman amino acid analyzer according to the manufacturer's instructions. Separate injections of AOrn and CP plus enzyme at concentrations identical to those used in the enzyme assay as well as Cit and ACit were used as controls and standards.

Amino acid analyses were performed by Crystal Stroud at the Clinical Lab at Children's National Medical Center, Washington DC.

Synthesis of compounds

Acetyl Citrulline synthesis

L-Citrulline (0.1mmols) was dissolved in 1ml of water and incubated with 0.2mmols of acetic anhydride for 2 hours at room temperature. After incubation the solution was neutralized by the addition of 1% (w/v) KOH to pH 7. Purity was assessed by LC-MS.

PALAO synthesis

PALO (5.1mg or 0.02mmols) was dissolved in 200µl of water and incubated with 0.2mmols of acetic anhydride for 2 hours at room temperature with constant stirring. After incubation the solution was neutralized with 2ml of 1% (w/v) KOH. Sample purity was assessed by LC-MS.

N-Succinyl-L-Orn synthesis

t-Boc-Ornithine (1.076mmols) was dissolved in 5ml of water. pH was kept neutral with 40% (w/v) NaOH while 1.076mmols of succinic anhydride was added. Once the pH was neutral the reaction was considered finalized. The Boc group was removed by acidification of the sample with trifluoroacetic acid and then neutralized. The purity of the sample was assessed by LC-MS.

PALSO synthesis

PALO (5.1mg or 0.02mmols) was dissolved in 200µl of water. pH was kept neutral with 40% (w/v) NaOH while 0.2mmols of succinic anhydride were added at room temperature with constant stirring. Once the pH stabilized the reaction was considered finalized and the purity of the sample was assessed by LC-MS.

N-Succinyl-L-Glutamate (NSG) and N-Succinyl-L-Aspartate (NSA) synthesis

Glutamic acid (monosodium salt, 0.025 mols) or aspartic acid (0.025 mols) was dissolved in 25ml of water. NaOH was added until fully dissolved. The solution was then incubated with 0.025 moles of succinic anhydride for 24 hours at around 50°C with constant stirring. Purity was assessed by LC-MS as described in Section 2.17. Quantification of the product formation was obtained by measuring the depletion of peaks corresponding to unreacted glutamic acid (148 m/z) using known concentrations of glutamic acid as a standard curve by LC-MS.

Bibliography

1. Glansdorff, N. and Y. Xu, *Microbial Arginine Biosynthesis: Pathway, Regulation and Industrial Production*. Microbiology Monographs, Amino Acid Biosynthesis ~ Pathways, Regulation and Metabolic Engineering, ed. S.B. Heidelberg. Vol. 5. 2006: Springer Berlin / Heidelberg. 219-257.
2. Lu, C.D., *Pathways and regulation of bacterial arginine metabolism and perspectives for obtaining arginine overproducing strains.* Appl Microbiol Biotechnol, 2006. **70**(3): p. 261-72.
3. Vogel, H.J., *Path of Ornithine Synthesis in Escherichia Coli.* Proc Natl Acad Sci U S A, 1953. **39**(7): p. 578-83.
4. Vogel, H.J., P.H. Abelson, and E.T. Bolton, *Short communications and preliminary notes on ornithine and proline synthesis in escherichia coli.* Biochim Biophys Acta, 1953. **11**(4): p. 584-5.
5. Glansdorff, N., G. Sand, and C. Verhoef, *The dual genetic control of ornithine transcarbamylase synthesis in Escherichia coli K12.* Mutat Res, 1967. **4**(6): p. 743-51.
6. Kwon, D.H., C.D. Lu, D.A. Walthall, et al., *Structure and regulation of the carAB operon in Pseudomonas aeruginosa and Pseudomonas stutzeri: no untranslated region exists.* J Bacteriol, 1994. **176**(9): p. 2532-42.
7. Mergeay, M., D. Gigot, J. Beckmann, et al., *Physiology and genetics of carbamoylphosphate synthesis in Escherichia coli K12.* Mol Gen Genet, 1974. **133**(4): p. 299-316.
8. Paulus, T.J. and R.L. Switzer, *Characterization of pyrimidine-repressible and arginine-repressible carbamyl phosphate synthetases from Bacillus subtilis.* J Bacteriol, 1979. **137**(1): p. 82-91.
9. Cunin, R., N. Glansdorff, A. Pierard, et al., *Biosynthesis and metabolism of arginine in bacteria.* Microbiol Rev, 1986. **50**(3): p. 314-52.
10. Marina, A., P.M. Alzari, J. Bravo, et al., *Carbamate kinase: New structural machinery for making carbamoyl phosphate, the common precursor of pyrimidines and arginine.* Protein Sci, 1999. **8**(4): p. 934-40.
11. Uriarte, M., A. Marina, S. Ramon-Maiques, et al., *The carbamoyl-phosphate synthetase of Pyrococcus furiosus is enzymologically and structurally a carbamate kinase.* J Biol Chem, 1999. **274**(23): p. 16295-303.
12. Udaka, S., Kinoshita, S., *Studies on L-ornithine fermentation; I. The biosynthetic pathway of L-ornithine in Micrococcus glutamicus.* J. Gen. Appl. Microbiol, 1958. **4**(4): p. 272-282.
13. Xu, Y., B. Labedan, and N. Glansdorff, *Surprising arginine biosynthesis: a reappraisal of the enzymology and evolution of the pathway in microorganisms.* Microbiol Mol Biol Rev, 2007. **71**(1): p. 36-47.
14. Hani, E.K., D. Ng, and V.L. Chan, *Arginine biosynthesis in Campylobacter jejuni TGH9011: determination of the argCOBD cluster.* Can J Microbiol, 1999. **45**(11): p. 959-69.
15. Xu, Y., N. Glansdorff, and B. Labedan, *Bioinformatic analysis of an unusual gene-enzyme relationship in the arginine biosynthetic pathway among marine gamma proteobacteria: implications concerning the formation of N-acetylated intermediates in prokaryotes.* BMC Genomics, 2006. **7**: p. 4.
16. Errey, J.C. and J.S. Blanchard, *Functional characterization of a novel ArgA from Mycobacterium tuberculosis.* J Bacteriol, 2005. **187**(9): p. 3039-44.
17. Abadjieva, A., K. Pauwels, P. Hilven, et al., *A new yeast metabolon involving at least the two first enzymes of arginine biosynthesis: acetylglutamate*

synthase activity requires complex formation with acetylglutamate kinase. J Biol Chem, 2001. **276**(46): p. 42869-80.

18. Parsot, C., A. Boyen, G.N. Cohen, et al., *Nucleotide sequence of Escherichia coli argB and argC genes: comparison of N-acetylglutamate kinase and N-acetylglutamate-gamma-semialdehyde dehydrogenase with homologous and analogous enzymes.* Gene, 1988. **68**(2): p. 275-83.
19. Leisinger, T. and D. Haas, *N-Acetylglutamate synthase of Escherichia coli regulation of synthesis and activity by arginine.* J Biol Chem, 1975. **250**(5): p. 1690-3.
20. Marc, F., P. Weigel, C. Legrain, et al., *Characterization and kinetic mechanism of mono- and bifunctional ornithine acetyltransferases from thermophilic microorganisms.* Eur J Biochem, 2000. **267**(16): p. 5217-26.
21. Sakanyan, V., D. Charlier, C. Legrain, et al., *Primary structure, partial purification and regulation of key enzymes of the acetyl cycle of arginine biosynthesis in Bacillus stearothermophilus: dual function of ornithine acetyltransferase.* J Gen Microbiol, 1993. **139**(3): p. 393-402.
22. Ramon-Maiques, S., M.L. Fernandez-Murga, F. Gil-Ortiz, et al., *Structural bases of feed-back control of arginine biosynthesis, revealed by the structures of two hexameric N-acetylglutamate kinases, from Thermotoga maritima and Pseudomonas aeruginosa.* J Mol Biol, 2006. **356**(3): p. 695-713.
23. Sancho-Vaello, E., M.L. Fernandez-Murga, and V. Rubio, *Site-directed mutagenesis studies of acetylglutamate synthase delineate the site for the arginine inhibitor.* FEBS Lett, 2008. **582**(7): p. 1081-6.
24. Sancho-Vaello, E., M.L. Fernandez-Murga, and V. Rubio, *Mechanism of arginine regulation of acetylglutamate synthase, the first enzyme of arginine synthesis.* FEBS Lett, 2009. **583**(1): p. 202-6.
25. Maheswaran, M., C. Urbanke, and K. Forchhammer, *Complex formation and catalytic activation by the PII signaling protein of N-acetyl-L-glutamate kinase from Synechococcus elongatus strain PCC 7942.* J Biol Chem, 2004. **279**(53): p. 55202-10.
26. Arcondeguy, T., R. Jack, and M. Merrick, *P(II) signal transduction proteins, pivotal players in microbial nitrogen control.* Microbiol Mol Biol Rev, 2001. **65**(1): p. 80-105.
27. Deibel, R.H., *Utilization of arginine as an energy source for the growth of Streptococcus faecalis.* J Bacteriol, 1964. **87**(5): p. 988-92.
28. Schimke, R.T., C.M. Berlin, E.W. Sweeney, et al., *The generation of energy by the arginine dihydrolase pathway in Mycoplasma hominis 07.* J Biol Chem, 1966. **241**(10): p. 2228-36.
29. Broman, K., N. Lauwers, V. Stalon, et al., *Oxygen and nitrate in utilization by Bacillus licheniformis of the arginase and arginine deiminase routes of arginine catabolism and other factors affecting their syntheses.* J Bacteriol, 1978. **135**(3): p. 920-7.
30. Vander Wauven, C., A. Pierard, M. Kley-Raymann, et al., *Pseudomonas aeruginosa mutants affected in anaerobic growth on arginine: evidence for a four-gene cluster encoding the arginine deiminase pathway.* J Bacteriol, 1984. **160**(3): p. 928-34.
31. Blakemore, R.P. and E. Canale-Parola, *Arginine catabolism by Treponema denticola.* J Bacteriol, 1976. **128**(2): p. 616-22.

32. Laishley, E.J. and R.W. Bernlohr, *The regulation and kinetics of the two ornithine transcarbamylase enzymes of Bacillus licheniformis.* Biochim Biophys Acta, 1968. **167**(3): p. 547-54.
33. Legrain, C., V. Stalon, J.P. Noullez, et al., *Structure and function of ornithine carbamoyltransferases.* Eur J Biochem, 1977. **80**(2): p. 401-9.
34. Ramos, F., V. Stalon, A. Pierard, et al., *The specialization of the two ornithine carbamoyltransferases of Pseudomonas.* Biochim Biophys Acta, 1967. **139**(1): p. 98-106.
35. Stalon, V., F. Ramos, A. Pierard, et al., *The occurrence of a catabolic and an anabolic ornithine carbamoyltransferase in Pseudomonas.* Biochim Biophys Acta, 1967. **139**(1): p. 91-7.
36. Itoh, Y., L. Soldati, V. Stalon, et al., *Anabolic ornithine carbamoyltransferase of Pseudomonas aeruginosa: nucleotide sequence and transcriptional control of the argF structural gene.* J Bacteriol, 1988. **170**(6): p. 2725-34.
37. Haas, D., B.W. Holloway, A. Schambock, et al., *The genetic organization of arginine biosynthesis in Pseudomonas aeruginosa.* Mol Gen Genet, 1977. **154**(1): p. 7-22.
38. Tricot, C., S. Schmid, H. Baur, et al., *Catabolic ornithine carbamoyltransferase of Pseudomonas aeruginosa. Changes of allosteric properties resulting from modifications at the C-terminus.* Eur J Biochem, 1994. **221**(1): p. 555-61.
39. Mira De Orduna, R., M.L. Patchett, S.Q. Liu, et al., *Growth and arginine metabolism of the wine lactic acid bacteria Lactobacillus buchneri and Oenococcus oeni at different pH values and arginine concentrations.* Appl Environ Microbiol, 2001. **67**(4): p. 1657-62.
40. Stalon, V., F. Ramos, A. Pierard, et al., *Regulation of the catabolic ornithine carbamoyltransferase of Pseudomonas fluorescens. A comparison with the anabolic transferase and with a mutationally modified catabolic transferase.* Eur J Biochem, 1972. **29**(1): p. 25-35.
41. Morizono, H., M. Tuchman, B.S. Rajagopal, et al., *Expression, purification and kinetic characterization of wild-type human ornithine transcarbamylase and a recurrent mutant that produces 'late onset' hyperammonaemia.* Biochem J, 1997. **322 (Pt 2)**: p. 625-31.
42. Labedan, B., A. Boyen, M. Baetens, et al., *The evolutionary history of carbamoyltransferases: A complex set of paralogous genes was already present in the last universal common ancestor.* J Mol Evol, 1999. **49**(4): p. 461-73.
43. Wild, J.R. and M.E. Wales, *Molecular evolution and genetic engineering of protein domains involving aspartate transcarbamoylase.* Annu Rev Microbiol, 1990. **44**: p. 193-218.
44. Kantrowitz, E.R. and W.N. Lipscomb, *Escherichia coli aspartate transcarbamoylase: the molecular basis for a concerted allosteric transition.* Trends Biochem Sci, 1990. **15**(2): p. 53-9.
45. Labedan, B., Y. Xu, D.G. Naumoff, et al., *Using quaternary structures to assess the evolutionary history of proteins: the case of the aspartate carbamoyltransferase.* Mol Biol Evol, 2004. **21**(2): p. 364-73.
46. Naumoff, D.G., Y. Xu, N. Glansdorff, et al., *Retrieving sequences of enzymes experimentally characterized but erroneously annotated : the case of the putrescine carbamoyltransferase.* BMC Genomics, 2004. **5**(1): p. 52.

47. Shi, D., R. Gallegos, J. DePonte, 3rd, et al., *Crystal structure of a transcarbamylase-like protein from the anaerobic bacterium Bacteroides fragilis at 2.0 A resolution.* J Mol Biol, 2002. **320**(4): p. 899-908.
48. Sekowska, A., A. Danchin, and J.L. Risler, *Phylogeny of related functions: the case of polyamine biosynthetic enzymes.* Microbiology, 2000. **146 (Pt 8)**: p. 1815-28.
49. Babbitt, P.C., *Definitions of enzyme function for the structural genomics era.* Curr Opin Chem Biol, 2003. **7**(2): p. 230-7.
50. Brenner, S.E., *Errors in genome annotation.* Trends Genet, 1999. **15**(4): p. 132-3.
51. Gerlt, J.A. and P.C. Babbitt, *Can sequence determine function?* Genome Biol, 2000. **1**(5): p. REVIEWS0005.
52. Shi, D., H. Morizono, X. Yu, et al., *Human ornithine transcarbamylase: crystallographic insights into substrate recognition and conformational changes.* Biochem J, 2001. **354**(Pt 3): p. 501-9.
53. Honzatko, R.B., J.L. Crawford, H.L. Monaco, et al., *Crystal and molecular structures of native and CTP-liganded aspartate carbamoyltransferase from Escherichia coli.* J Mol Biol, 1982. **160**(2): p. 219-63.
54. Jin, L., B.A. Seaton, and J.F. Head, *Crystal structure at 2.8 A resolution of anabolic ornithine transcarbamylase from Escherichia coli.* Nat Struct Biol, 1997. **4**(8): p. 622-5.
55. Murata, L.B. and H.K. Schachman, *Structural similarity between ornithine and aspartate transcarbamoylases of Escherichia coli: characterization of the active site and evidence for an interdomain carboxy-terminal helix in ornithine transcarbamoylase.* Protein Sci, 1996. **5**(4): p. 709-18.
56. Allewell, N.M., D. Shi, H. Morizono, et al., *Molecular Recognition by Ornithine and Aspartate Transcarbamylases.* Accounts of Chemical Research, 1999. **32**(10): p. 885-894.
57. Houghton, J.E., D.A. Bencini, G.A. O'Donovan, et al., *Protein differentiation: a comparison of aspartate transcarbamoylase and ornithine transcarbamoylase from Escherichia coli K-12.* Proc Natl Acad Sci U S A, 1984. **81**(15): p. 4864-8.
58. Van Vliet, F., R. Cunin, A. Jacobs, et al., *Evolutionary divergence of genes for ornithine and aspartate carbamoyl-transferases--complete sequence and mode of regulation of the Escherichia coli argF gene; comparison of argF with argI and pyrB.* Nucleic Acids Res, 1984. **12**(15): p. 6277-89.
59. Naumoff, D.G., Y. Xu, V. Stalon, et al., *The difficulty of annotating genes: the case of putrescine carbamoyltransferase.* Microbiology, 2004. **150**(Pt 12): p. 3908-11.
60. Crooks, G.E., G. Hon, J.M. Chandonia, et al., *WebLogo: a sequence logo generator.* Genome Res, 2004. **14**(6): p. 1188-90.
61. Morizono, H., J. Cabrera-Luque, D. Shi, et al., *Acetylornithine transcarbamylase: a novel enzyme in arginine biosynthesis.* J Bacteriol, 2006. **188**(8): p. 2974-82.
62. Assis, M.P., L.R. Mariano, S.J. Michereff, et al., *Antagonism of yeasts to Xanthomonas campestris pv. campestris on cabbage phylloplane in field.* Revista de Microbiologia, 1999. **30**: p. 191-195.
63. Shi, D., H. Morizono, X. Yu, et al., *Crystal structure of N-acetylornithine transcarbamylase from Xanthomonas campestris: a novel enzyme in a new*

arginine biosynthetic pathway found in several eubacteria. J Biol Chem, 2005. **280**(15): p. 14366-9.
64. Pastra-Landis, S.C., J. Foote, and E.R. Kantrowitz, *An improved colorimetric assay for aspartate and ornithine transcarbamylases.* Anal Biochem, 1981. **118**(2): p. 358-63.
65. Strandholm, J.J., N.R. Buist, N.G. Kennaway, et al., *Excretion of -N-acetylcitrulline in citrullinaemia.* Biochim Biophys Acta, 1971. **244**(1): p. 214-6.
66. Clantin, B., C. Tricot, T. Lonhienne, et al., *Probing the role of oligomerization in the high thermal stability of Pyrococcus furiosus ornithine carbamoyltransferase by site-specific mutants.* Eur J Biochem, 2001. **268**(14): p. 3937-42.
67. Selwyn, M.J., *A simple test for inactivation of an enzyme during assay.* Biochim Biophys Acta, 1965. **105**(1): p. 193-5.
68. Simon, J.P. and V. Stalon, *L-Ornithine carbamoyltransferase from Saccharomyces cerevisiae: steady-state kinetic analysis.* Eur J Biochem, 1977. **75**(2): p. 571-81.
69. Legrain, C. and V. Stalon, *Ornithine carbamoyltransferase from Escherichia coli W. Purification, structure and steady-state kinetic analysis.* Eur J Biochem, 1976. **63**(1): p. 289-301.
70. Bates, M., R.L. Weiss, and S. Clarke, *Ornithine transcarbamylase from Neurospora crassa: purification and properties.* Arch Biochem Biophys, 1985. **239**(1): p. 172-83.
71. Baur, H., C. Tricot, V. Stalon, et al., *Converting catabolic ornithine carbamoyltransferase to an anabolic enzyme.* J Biol Chem, 1990. **265**(25): p. 14728-31.
72. Kuo, L.C., W. Herzberg, and W.N. Lipscomb, *Substrate specificity and protonation state of ornithine transcarbamoylase as determined by pH studies.* Biochemistry, 1985. **24**(18): p. 4754-61.
73. LiCata, V.J. and N.M. Allewell, *Is substrate inhibition a consequence of allostery in aspartate transcarbamylase?* Biophys Chem, 1997. **64**(1-3): p. 225-34.
74. Vissers, S., L.M. Urrestarazu, J.C. Jauniaux, et al., *Inhibition of ornithine carbamoyltransferase by arginase among Yeast: correlation with energy production, subcellular localization and enzyme synthesis.* J. Gen. Microbiol, 1982. **128**: p. 1235-1247.
75. Penninckx, M. and D. Gigot, *Synthesis and interaction with Escherichia coli L-ornithine carbamolytransferase of two potential transition-state analogues.* FEBS Lett, 1978. **88**(1): p. 94-6.
76. Haas, D. and T. Leisinger, *Multiple control of N-acetylglutamate synthetase from Pseudomonas aeruginosa: synergistic inhibition by acetylglutamate and polyamines.* Biochem Biophys Res Commun, 1974. **60**(1): p. 42-7.
77. Tricot, C., V. Villeret, G. Sainz, et al., *Allosteric regulation in Pseudomonas aeruginosa catabolic ornithine carbamoyltransferase revisited: association of concerted homotropic cooperative interactions and local heterotropic effects.* J Mol Biol, 1998. **283**(3): p. 695-704.
78. Wexler, H.M., *Bacteroides: the good, the bad, and the nitty-gritty.* Clin Microbiol Rev, 2007. **20**(4): p. 593-621.

79. Backhed, F., R.E. Ley, J.L. Sonnenburg, et al., *Host-bacterial mutualism in the human intestine.* Science, 2005. **307**(5717): p. 1915-20.
80. Polk, B.F. and D.L. Kasper, *Bacteroides fragilis subspecies in clinical isolates.* Ann Intern Med, 1977. **86**(5): p. 569-71.
81. Caldovic, L., H. Morizono, X. Yu, et al., *Identification, cloning and expression of the mouse N-acetylglutamate synthase gene.* Biochem J, 2002. **364**(Pt 3): p. 825-31.
82. Qu, Q., H. Morizono, D. Shi, et al., *A novel bifunctional N-acetylglutamate synthase-kinase from Xanthomonas campestris that is closely related to mammalian N-acetylglutamate synthase.* BMC Biochem, 2007. **8**: p. 4.
83. Marco-Marin, C., S. Ramon-Maiques, S. Tavarez, et al., *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, 2003. **334**(3): p. 459-76.
84. Haas, D. and T. Leisinger, *N-acetylglutamate 5-phosphotransferase of Pseudomonas aeruginosa. Catalytic and regulatory properties.* Eur J Biochem, 1975. **52**(2): p. 377-93.
85. Shi, D., H. Morizono, J. Cabrera-Luque, et al., *Structure and catalytic mechanism of a novel N-succinyl-L-ornithine transcarbamylase in arginine biosynthesis of Bacteroides fragilis.* J Biol Chem, 2006. **281**(29): p. 20623-31.
86. Shi, D., X. Yu, L. Roth, et al., *Structures of N-acetylornithine transcarbamoylase from Xanthomonas campestris complexed with substrates and substrate analogs imply mechanisms for substrate binding and catalysis.* Proteins, 2006. **64**(2): p. 532-42.
87. Porter, R.W., M.O. Modebe, and G.R. Stark, *Aspartate transcarbamylase. Kinetic studies of the catalytic subunit.* J Biol Chem, 1969. **244**(7): p. 1846-59.
88. Wang, J., K.A. Stieglitz, J.P. Cardia, et al., *Structural basis for ordered substrate binding and cooperativity in aspartate transcarbamoylase.* Proc Natl Acad Sci U S A, 2005. **102**(25): p. 8881-6.
89. Lipscomb, W.N., *Aspartate transcarbamylase from Escherichia coli: activity and regulation.* Adv Enzymol Relat Areas Mol Biol, 1994. **68**: p. 67-151.
90. Shi, D., H. Morizono, M. Aoyagi, et al., *Crystal structure of human ornithine transcarbamylase complexed with carbamoyl phosphate and L-norvaline at 1.9 A resolution.* Proteins, 2000. **39**(4): p. 271-7.
91. Ha, Y., M.T. McCann, M. Tuchman, et al., *Substrate-induced conformational change in a trimeric ornithine transcarbamoylase.* Proc Natl Acad Sci U S A, 1997. **94**(18): p. 9550-5.
92. Shi, D., H. Morizono, Y. Ha, et al., *1.85-A resolution crystal structure of human ornithine transcarbamoylase complexed with N-phosphonacetyl-L-ornithine. Catalytic mechanism and correlation with inherited deficiency.* J Biol Chem, 1998. **273**(51): p. 34247-54.
93. Krupyanko, V.I., *Corrected equations for calculation of constants in enzyme inhibition and activation.* Biochemistry (Mosc), 2007. **72**(4): p. 380-91.
94. Visca, P., G. Colotti, L. Serino, et al., *Metal regulation of siderophore synthesis in Pseudomonas aeruginosa and functional effects of siderophore-metal complexes.* Appl Environ Microbiol, 1992. **58**(9): p. 2886-93.

95. Pandey, A. and R.V. Sonti, *Role of the FeoB protein and siderophore in promoting virulence of Xanthomonas oryzae pv. oryzae on rice.* J Bacteriol. **192**(12): p. 3187-203.
96. Schneider, B.L., A.K. Kiupakis, and L.J. Reitzer, *Arginine catabolism and the arginine succinyltransferase pathway in Escherichia coli.* J Bacteriol, 1998. **180**(16): p. 4278-86.
97. Vander Wauven, C. and V. Stalon, *Occurrence of succinyl derivatives in the catabolism of arginine in Pseudomonas cepacia.* J Bacteriol, 1985. **164**(2): p. 882-6.
98. Endrizzi, J.A., P.T. Beernink, T. Alber, et al., *Binding of bisubstrate analog promotes large structural changes in the unregulated catalytic trimer of aspartate transcarbamoylase: implications for allosteric regulation.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5077-82.
99. Shi, D., X. Yu, J. Cabrera-Luque, et al., *A single mutation in the active site swaps the substrate specificity of N-acetyl-L-ornithine transcarbamylase and N-succinyl-L-ornithine transcarbamylase.* Protein Sci, 2007. **16**(8): p. 1689-99.
100. Ramon-Maiques, S., A. Marina, F. Gil-Ortiz, et al., *Structure of acetylglutamate kinase, a key enzyme for arginine biosynthesis and a prototype for the amino acid kinase enzyme family, during catalysis.* Structure, 2002. **10**(3): p. 329-42.
101. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.
102. Caldovic, L., H. Morizono, M. Gracia Panglao, et al., *Cloning and expression of the human N-acetylglutamate synthase gene.* Biochem Biophys Res Commun, 2002. **299**(4): p. 581-6.

Appendix 1: Resumen en español

Discovery of novel pathways of microbial arginine biosynthesis

Thesis by

Juan Manuel Cabrera Luque

CHILDREN'S NATIONAL MEDICAL CENTER
Research Center for Genetic Medicine
Washington DC, USA
2010

Resumen en Español

Introducción

El aminoácido L-arginina es un componente esencial de todos los organismos vivos. Su importancia reside en la gran variedad de funciones que la arginina, así como los intermediarios involucrados en su síntesis, tienen en las células, como muestra la Figura 1.

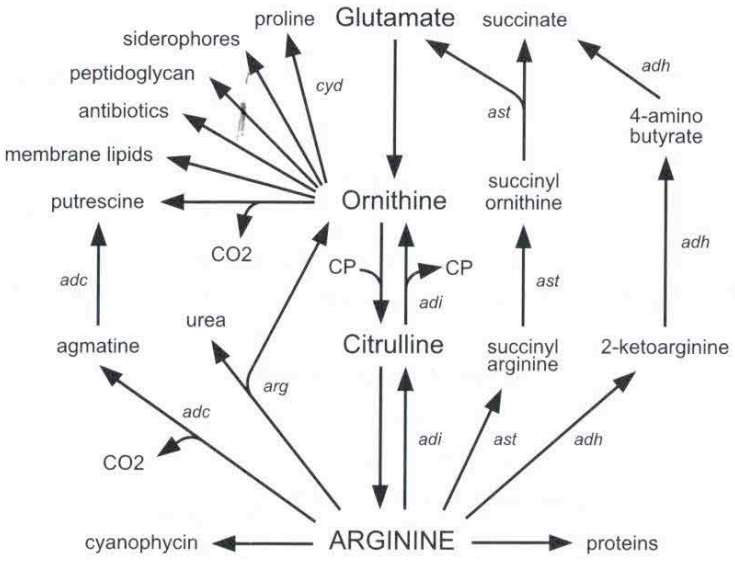


Figura 1: Representación esquemática de las principales rutas metabólicas interconectadas con la ruta de biosíntesis de la arginina. Sólo se muestran los intermediarios más significativos que interconectan las rutas.

Entre las principales funciones de la arginina en la célula se encuentran:

- forma parte de la composición aminoacídica de muchas proteínas
- precursor de los siguientes compuestos que almacenan energía
  - fosfoarginina, usado en la contracción muscular en invertebrados y
  - fosfocreatina, que se encuentra almacenado en el músculo y el cerebro en vertebrados
- metabolito intermediario en el ciclo de la urea
- precursor de las moléculas de señalización: oxido nítrico, agmatina y glutamato en mamíferos
- y es usada como compuesto de almacenamiento de nitrógeno en plantas y cianobacterias

En numerosos procariotas, hongos y plantas la síntesis *de novo* de la arginina ocurre a partir de glutamato en ocho pasos enzimáticos como se describe en la Figura 2. La acetilación de los precursores iniciales en la biosíntesis de la arginina permite la separación de ésta ruta metabólica de la biosíntesis de la prolina.

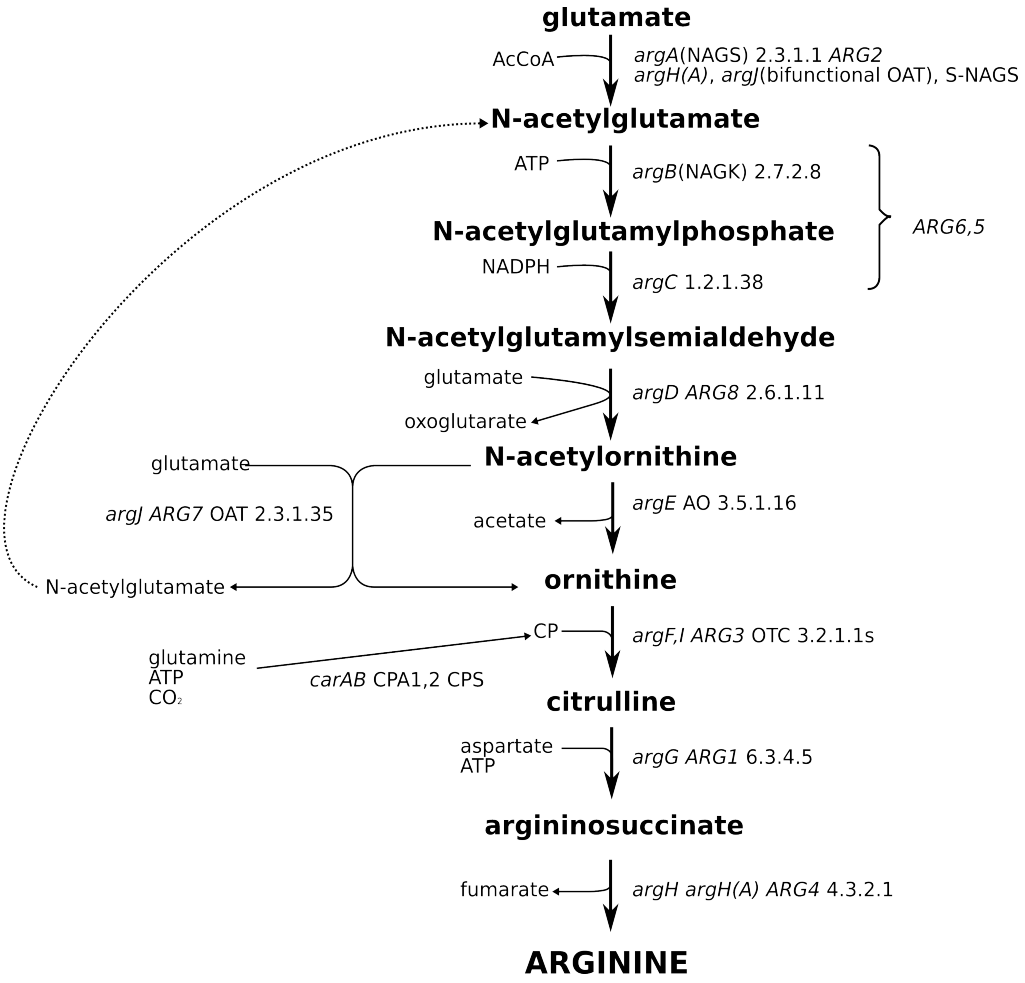


Figura 2: Esquema representativo de la ruta de biosíntesis de novo de la arginina a partir de L-glutamato.

El primer paso en la ruta de biosíntesis de la arginina es la acetilación del grupo α-amino del glutamato. Ésta reacción puede ser catalizada por varias enzimas en los procariotas:

(i) NAGS clásica, inicialmente caracterizada en *E. coli* y presente en los animales capaces de sintetizar arginina *de novo*.

(ii) OAT bifuncional, capaz de utilizar AcCoA o AOrn para acetilar el glutamato. Ésta enzima se encuentra en los reinos Bacteria, Archaea y Eukaryotes.

(iii) NAGS truncada, perteneciente a la familia de las GNAT, ya sea de manera independiente o bien fusionada con la argininosuccinato liasa.

En el caso de los organismos que utilizan la acetilornitina deacetilasa, la acetilación del glutamato es, aparentemente, el punto de control en el flujo de intermediarios acetilados que conducen a la ornitina. Se dice de ellos que poseen una ruta de biosíntesis de la arginina linear. La actividad de la NAGS clásica está controlada por un mecanismo de inhibición por retroalimentación ejercido por el producto final de la ruta: arginina. Lo mismo ocurre en el caso de los organismos que

utilizan la NAGS truncada. El mecanismo por el cual la actividad de las OATs, tanto mono como bifuncionales es controlada por la arginina no está demostrado, sin embargo éstas enzimas son inhibidas competitivamente por la ornitina, controlando así su actividad.

El segundo paso en la ruta de biosíntesis de la arginina es la fosforilación del N-acetilglutamato catalizada por la enzima NAGK (Figura 2). Al contrario que ocurre con la acetilación del glutamato, la fosforilación del N-acetilglutamato sólo es catalizada por NAGK, indicando la importancia que esta enzima tiene en el control de la ruta biosintética de la arginina.

En organismos tales como *E. coli* y *Micobacterium lacticum* entre otros, NAGK es insensible a la inhibición por arginina y por lo tanto la reacción catalizada por NAGS es la más relevante en el control de la ruta. Sin embargo en aquellos organismos en los que la acetilación del glutamato ocurre a partir de la AOrn, y por lo tanto poseen una ruta denominada cíclica, la reacción catalizada por NAGK es crucial en el control del flujo de metabolitos en la ruta de biosíntesis de la arginina.

De todos los pasos implicados en la ruta biosintética de la arginina, la acetilación de glutamato y su fosforilación han sido los más estudiados dada su importancia en la regulación del flujo de metabolitos en ésta ruta.

Sin embargo, la transcarbamilación de la ornitina constituye otra reacción crucial en la ruta de biosíntesis y biodegradación de la arginina como muestra la Figura 3.

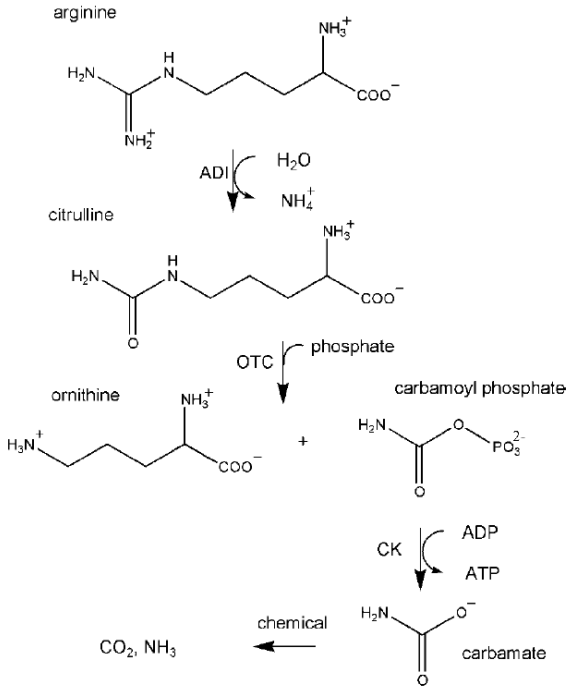


Figura 3: Representación esquemática de la ruta de degradación de la arginina por la arginina deiminasa.

En la ruta de biosíntesis, la ornitina es transcarbamilada por la enzima ornitina transcarbamilasa (OTC, EC 3.2.1.1). OTC pertenece a la familia de las transcarbamilasas cuyos miembros son identificados fácilmente basándose

únicamente en las homologías a nivel de las secuencias peptídicas que presentan todos ellos. Todos comparten el motivo de unión al fosfato de carbamilo, que es un sustrato común a todos los miembros de la familia. Los miembros más estudiados de la familia de las transcarbamilasas son OTC y aspartato transcarbamilasa (ATC).

Los estudios sobre la evolución de las transcarbamilasas realizados por Labedan y colaboradores revelaron la presencia de ésta actividad en el último ancestro común universal (Last Universal Common Ancestor, LUCA), destacando la importancia de la reacciones catalizadas por las transcarbamilasas.

Más recientemente, Naumoff y colaboradores, entre otros, han puesto de manifiesto cuan inexacta es la asignación de ciertas actividades enzimáticas basándose exclusivamente en las similitudes entre nuevas secuencias genómicas y las estudiadas bioquímicamente previamente. El trabajo de Naumoff y colaboradores se centró en el caso particular de las transcarbamilasas de putrescina (PTC), miembros de la familia de las transcarbamilasas incorrectamente anotadas como OTC. La topología del árbol evolutivo generado por Naumoff y colaboradores usando las 245 secuencias disponibles en el momento (Figura 4) no difiere de la del árbol presentado por Labedan y colaboradores anteriormente. En su análisis, Naumoff y colaboradores introdujeron dos nuevos grupos de transcarbamilasas (PTC y transcarbamilasas de función desconocida o UTC). El grupo de las UTC está formado por las transcarbamilasas anotadas como OTCasas pero que carecen de la capacidad de usar la ornitina como sustrato, como es el caso de la enzima, potencialmente transcarbamilasa, presente en el organismo *Bacteroides fragilis*.

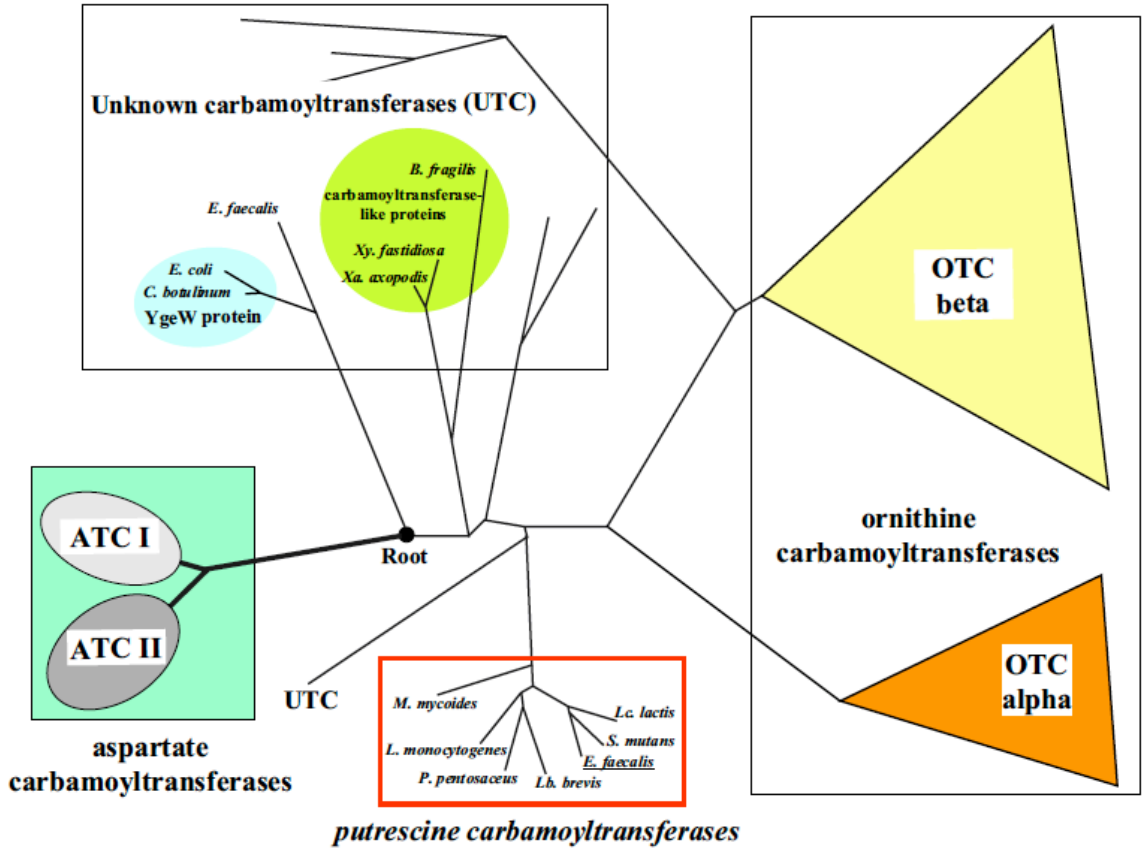


Figura 4: Esquema del árbol evolutivo de las transcarbamilasas. Adaptado de Naumoff et al. 2004.

El Dr. Malamy, usando mutagenesis mediada por transposon, demostró que en *B. fragilis* el gen que codifica para la posible ornitina transcarbamilasa está directamente involucrado en la biosíntesis de la arginina dado que cuando dicho gen no estaba presente en el genoma, el organismo se volvía auxotrofo para la arginina.

El gen *Bf*argF´ fue clonado y la proteína purificada y cristalizada por el Dr. Shi (Figura 5). La estructura tridimensional de la enzima presenta grandes similitudes con las estructuras de OTC y ATC, desafortunadamente la actividad de esta enzima no pudo ser identificada dado que no catalizaba la transcarbamilación de la ornitina.

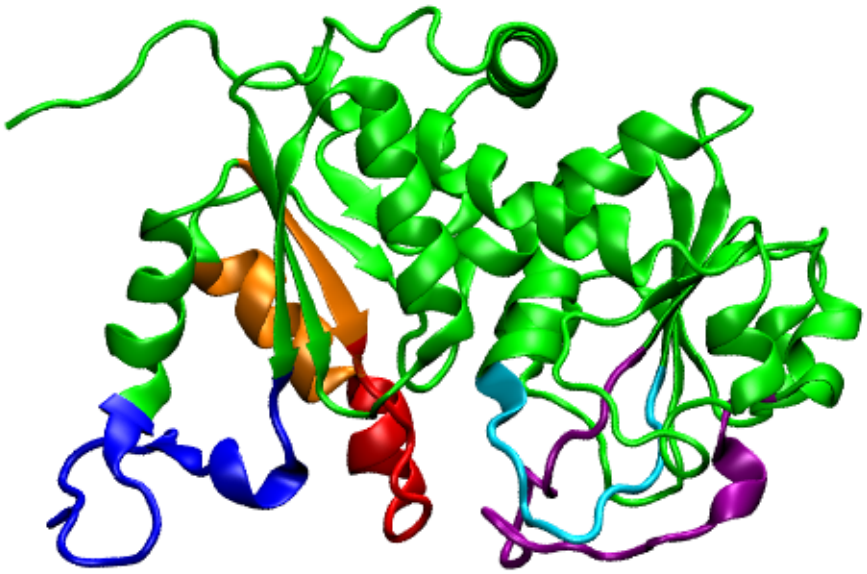


Figura 5: Representación de la estructura tridimensional del monómero the proteína similar a una transcarbamilasa presente en *B. fragilis*.

Usando como referencia la secuencia aminoacídica de la proteína de *B. fragilis* se pudieron identificar proteínas similares a ésta en los genomas de varios organismos entre los que se encuentran *Xylella fastidiosa*, *Cytophaga hutchinsonii*, *Xanthomonas campestris*, *Prevotella ruminicola* y *Tannerella forsythus*. Todas estas proteínas identificadas compartían con la de *B. fragilis* la ausencia del motivo de unión a ornitina (SMG), presente en todas las OTCasas. Asímismo, ninguno de los genes identificados en estos organismos pudo complementar los mutantes de *E. coli* (argF$^-$, argI$^-$) que carecen de OTCasa.

La primera actividad que se descubrió de entre las anteriormente mencionadas UTC fue la de *X. campestris* ArgF´. Dr. Morizono descubrió que la actividad de *Xc*ArgF´ era la de una nueva transcarbamilasa: acetilornitina transcarbamilasa (AOTCasa).

La presente tesis se centra en la caracterización bioquímica de la actividad de la enzima *Xc*ArgF´ así como en la identificación y caracterización de la actividad catalítica de la enzima *Bf*ArgF´ y en las implicaciones que la presencia de estas dos

enzimas tienen para la ruta de biosíntesis de la arginina en los organismos que las poseen.

A continuación presento el sumario de los resultados obtenidos durante la realización del presente proyecto de tesis: "Discovery of novel pathways of microbial arginine biosynthesis".

Objetivos de la presente tesis doctoral

1.- Caracterización bioquímica de la enzima ArgF' de *Xanthomonas campestris*.

- Se investigará la especificidad de la reacción catalizada por $_{Xc}$ArgF'.
- Los productos de la reacción catalizada por $_{Xc}$ArgF' serán identificados.
- Los parámetros cinéticos que caracterizan la reacción catalizada por $_{Xc}$ArgF' será determinados.
- La cinética de la reacción catalizada por $_{Xc}$ArgF' será estudiada en presencia de distintos inhibidores específicos.

2.- Diseño y síntesis de posibles sustratos así como de posibles compuestos inhibidores.

- Se identificarán los posibles sustratos de la enzima ArgF' de *B. fragilis* a partir de la estructura cuaternaria de la proteína. Los posibles sustratos de dicha enzima serán sintetizados así como los posibles sustratos de algunas de las otras enzimas implicadas en la biosíntesis de novo de la arginina en estos organismos.
- Así mismo se sintetizarán los compuestos con posible acción inhibitoria basándonos en la especificidad de la actividad enzimática de las proteínas ArgF' de *X. campestris* y *B. fragilis*.

3.- Se identificará la actividad de la enzima codificada por el gen argF' en *B. fragilis*. Dicha actividad será parcialmente caracterizada.

- La actividad enzimática de la proteína ArgF' de *B. fragilis* será identificada.
- La cinética de la reacción catalizada por la enzima ArgF' de *B. fragilis* será parcialmente caracterizada.

4.- Se determinarán las implicaciones que la presencia de estas transcarbamilasas tienen en la ruta metabólica de la biosíntesis *de novo* de la arginina en los organismos que las poseen.

- Se investigará el curso lógico de los metabolitos intermediarios en la biosíntesis de la arginina en *X. campestris*.
- Se investigarán los pasos iniciales de la ruta biosintética de la arginina en B. *fragilis* que conducen al sustrato de la enzima ArgF' en éste organismo.

Capítulo 2: "Functional characterization of a novel type of transcarbmylase that uses acetylornithine, identified in *Xanthomonas campestris pv campestris*"

Xanthomonas es un género que incluye varios patógenos de plantas. *X. campestris pv campestris* es una bacteria Gram-negativa que ataca a una gran variedad de crucíferas importantes para la economía. Entre las especies de crucíferas que *X. campestris pv campestris* es capaz de infectar se encuentran *Brassica* y *Arabidopsis* causando la enfermedad de la raíz negra.

Prácticamente todos los genes involucrados en la biosíntesis de la arginina han sido identificados en *X. campestris* y *X. axonopodis* en base a las homologías que las secuencias de dichos genes presentan con los de otros organismos previamente secuenciados. No obstante, la secuencia de los genes que codifican para la posible OTCasa en estos organismos (argF'), son significativamente diferentes de la considerada secuencia canónica para ésta enzima.

Al comienzo de este proyecto se había identificado la posible actividad de éstas enzimas como la de un nuevo miembro de las transcarbamilasas: N-acetilornitina transcarbamilasa (AOTCasa).

En este capítulo se presentan pruebas concluyentes demostrando que la actividad catalizada por la enzima $_{Xc}$ArgF´ es la de una acetilornitina transcarbamilasa (AOTCasa) mediante la identificación tanto de los sustratos como de los productos de la reacción. Dichas pruebas fueron obtenidas por diferentes métodos entre los que se incluyen:

- La capacidad del producto de la reacción para reaccionar con la ninhidrina (Apartado 2.2.1 de la tesis)
- Análisis por LC-MS y NMR del producto de la reacción (Apartado 2.2.2 de la tesis)

La especificidad de la reacción catalizada por la $_{Xc}$AOTCasa ha sido demostrada indicando que acetilornitina es el único compuesto que junto con carbamilfosfato puede ser transcarbamilado para formar acetilcitrulina (Figura 6).

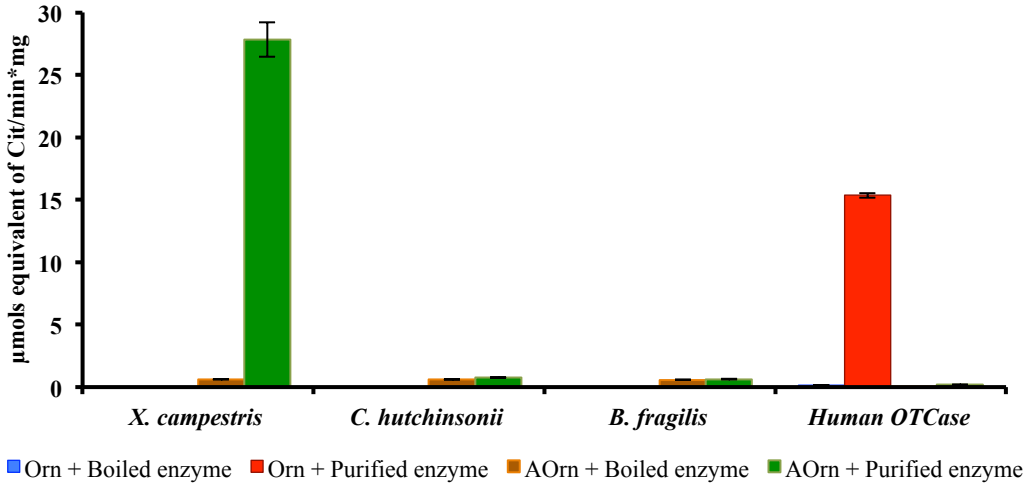


Figura 6: Comparación de la habilidad de diferentes ArgF' para utilizar ornitina o acetilornitina como sustrato.

Asimismo se han optimizado las condiciones para el ensayo de esta nueva actividad AOTCasa (Apartado 2.3.3 de la tesis).

Los parámetros cinéticos (K_m y V_{max}) han sido determinados para ambos sustratos (AOrn y CP, Tabla 1).

Tabla 1: Resumen de los parámetros cinéticos calculados para ambos sustratos de la reacción catalizada por XcArgF'.

Substrate	Km (mM)	Vmax (µmols/min*µg)
CP	0.01 ± 0.0007	50.68 ± 0.48
AOrn	1.05 ± 0.23	65.05 ± 4.77

Al contrario que ocurre con otros miembros de la familia de las transcarbamilasas, *Xc*AOTCasa no está sujeta a inhibición por el sustrato de la reacción (AOrn), incluso a concentraciones tan elevadas como 50mM (Figura 7).

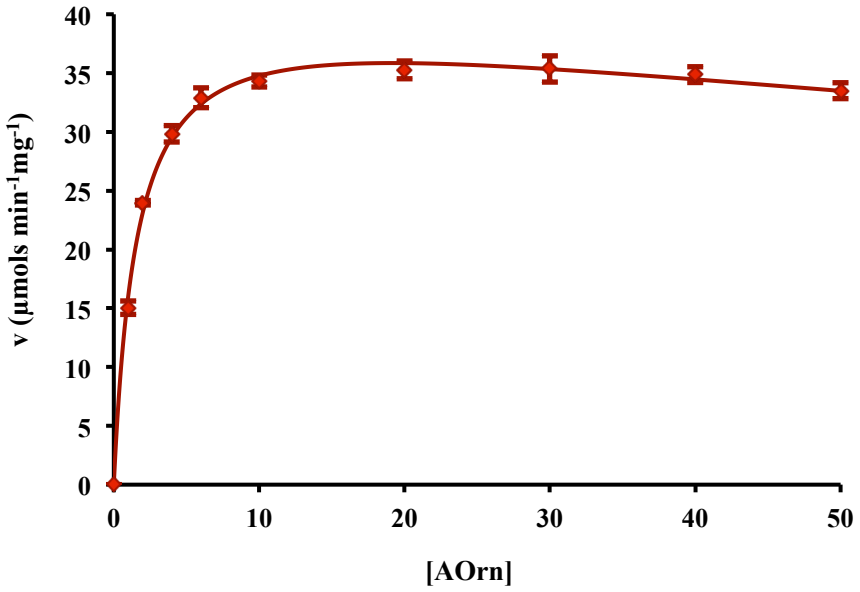


Figura 7: Estudio de la inhibición por el sustrato AOrn de la actividad de la *Xc*AOTCasa.

Estos resultados parecen indicar que la reacción catalizada por *Xc*AOTCasa no es crucial en el control de la ruta biosintética de la arginina, sin embargo el echo de que la ausencia de este gen confiera auxotrofía para la arginina hace resaltar su importancia en la mencionada ruta de biosíntesis.

Asimismo en este capítulo se demuestra que la acetilación del grupo alfa-amino no sólo es necesaria para el reconocimiento del sustrato sino que también es un requerimiento para los inhibidores, como muestran los resultados obtenidos en los estudios de inhibición por norvalina y acetilnorvalina (apartado 2.4.2 de la tesis), así como en los estudios con N^{δ}-fosfoacetil-L-ornitina (PALO) y N^{δ}-fosfoacetil-L-acetilornitina (PALAO), Figura 8 (apartado 2.4.3 de la tesis).

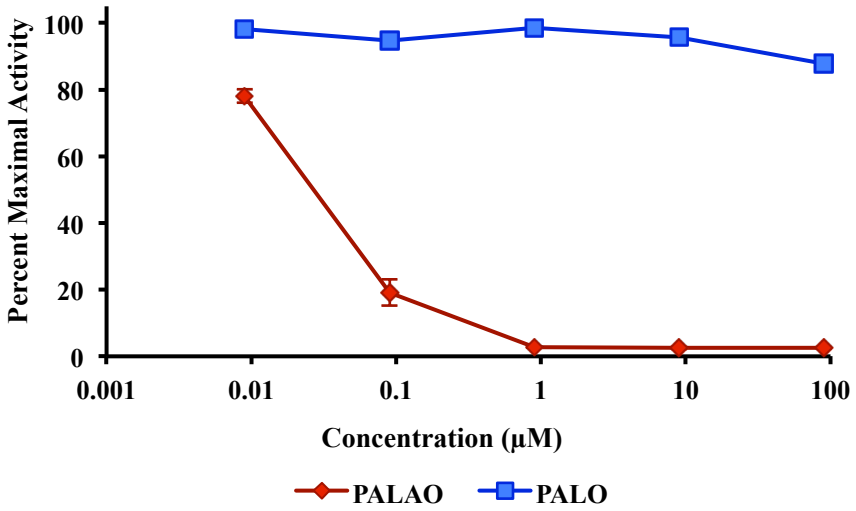


Figura 8: Estudios de inhibición de la actividad de XcAOTCasa por PALO y PALAO (sintetizado químicamente y usado sin posterior purificación).

El descubrimiento de ésta nueva actividad enzimática (AOTCasa), implica que el producto de la reaccíon catalizada por *Xc*ArgF' (acetilcitrulina) ha de continuar la ruta metabólica para sintetizar arginina. De ser cierto la siguiente enzima de la ruta debería ser capaz de deacetilar la acetilcitrullina.

La capacidad de la enzima ArgE de X. campestris para deacetilar acetilcitrulina y acetilornitina fue estudiada (apartado 2.5 de la tesis) mediante el uso de un analizador de aminoácidos Figura 9.

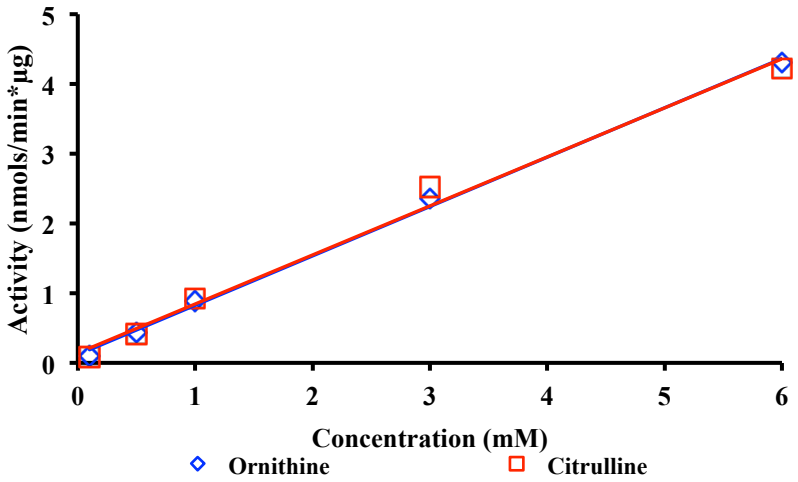


Figura 9: Estudio de la capacidad para deacetilar acetilcitrulina y acetilornitina de *Xc*ArgE.

Estos estudios han demostrado que la ruta de biosíntesis *de novo* de la arginina en los organismos que poseen una AOTCasa en lugar de una OTCasa difiere de la hasta ahora considerada canónica (Figura 2).

Basándonos en los resultados expuestos en este capítulo de la presente tesis proponemos la siguiente (Figura 10) ruta de biosíntesis *de novo* de la arginina en los organismos dotados de AOTCasa:

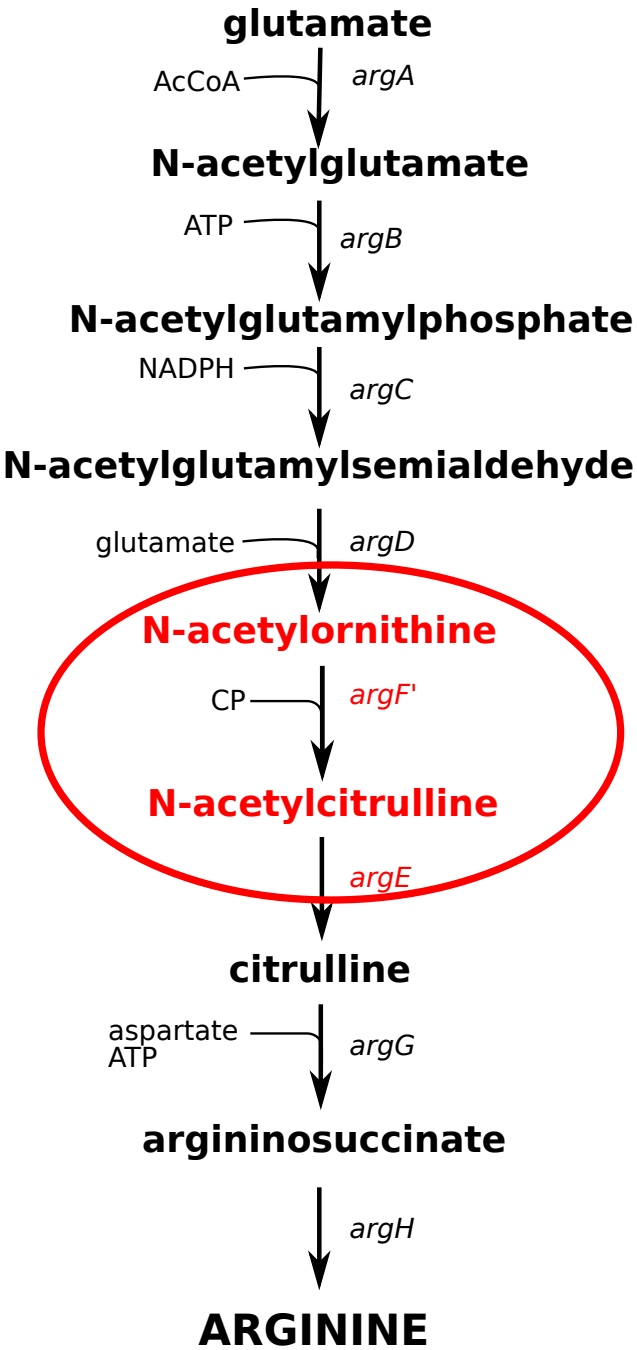


Figura 10: Representación esquemática de la ruta propuesta para la biosítesis *de novo* de la arginina en organismos dotados de AOTCasa.

Capítulo 3: "Demonstration that a structurally obvious trans carbamylase from *Bacteroides fragilis*, of previously unknown activity catalyzes the transcarbamylation of succinylornithine. Funtional characterization of this transcarbamylase"

En este capitulo se presentan los resultados que prueban que la actividad enzimática de la proteína codificada por el gen *argF´* de *B. fragilis* es la de una succinilornitina transcarbamilasa (SOTCasa).

Para ello se han identificado los posibles sustratos de la enzima basándonos en la información obtenida de la estructura cristalográfica de la proteína (apartado 3.2 de la tesis) y se han sintetizado químicamente dichos sustratos (apartado 3.3 de la tesis).

El primer compuesto sintetizado y ensayado fue la succinilornitina (Figura 11).

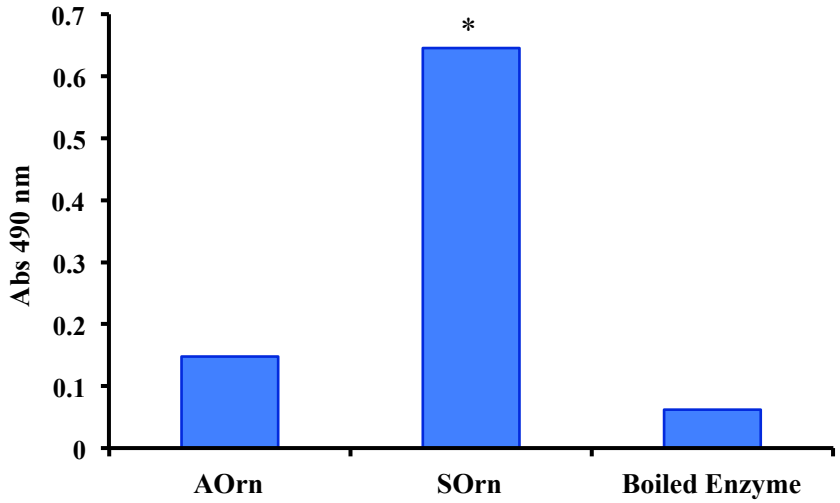


Figura 11: Primer ensayo para elucidar la naturaleza del segundo sustrato de la proteína *Bf*ArgF'.

Aparentemente la succinilornitina podría ser el sustrato de la enzima *Bf*ArgF'. Los análisis por resonacia magnética nuclear (NMR) y cromatografía liquida acoplada a un espectrómetro de masas (LC-MS) confirmaron que el producto de la reacción catalizada por *Bf*SOTCasa en presencia de carbamilfosfato y succinilornitina es ls succinilcitrullina (apartado 3.4 de la tesis).

Asimismo se presentan los resultados de la caracterización inicial de la enzima *Bf*SOTCasa y se han determinado los parámetros cinéticos para ambos sustratos de la reacción (Tabla 2).

Tabla 2: Resumen de los parámetros cinéticos calculados para ambos sustratos de la reacción catalizada por *Bf*ArgF'.

	SOrn	CP
K_m (mM)	12.5 ± 1.5	0.025 ± 0.003
V_{max} (µmol $min^{-1}mg^{-1}$)	370.6 ± 32.6	141.7 ± 3
K_{si} (mM)	20.4 ± 3.3	n/a

Al contrario que la *Xc*AOTCasa, *Bf*SOTCasa muestra inhibición por el sustrato (Figura 12) aunque dicha inhibición es mucho menos pronunciada comparada con otros miembros de la familia de las transcarbamilasas.

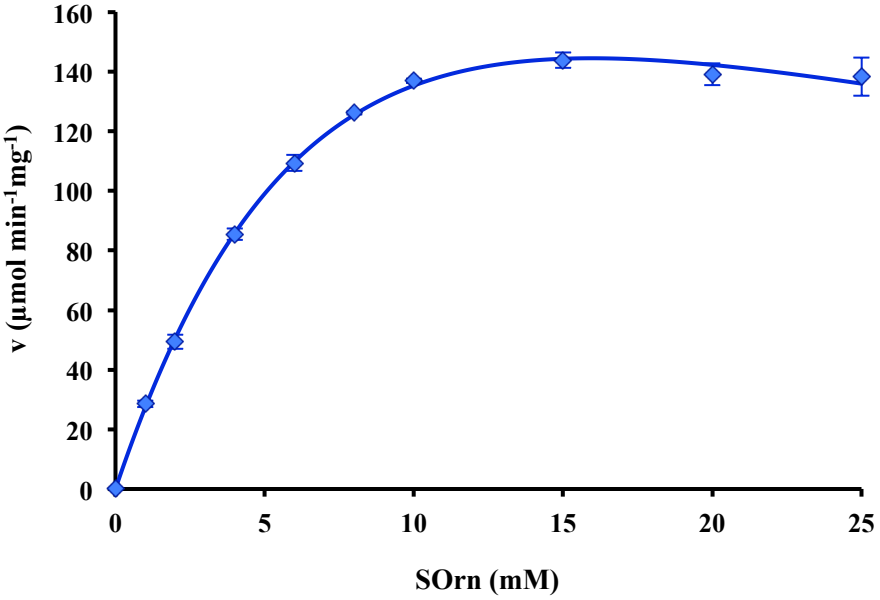


Figura 12: Curva de saturación por succinilornitina de la *Bf*SOTCasa.

Al igual que ocurría con *Xc*AOTCasa, la *Bf*SOTCasa requiere la presencia del grupo succinilo unido al extremo alfa-amino tanto para el reconocimiento del sustrato (succinilornitina), como de los inhibidores. Ésto ha sido demostrado mediante estudios de inhibición por PALO y PALSO (apartado 3.7 de la tesis).

El descubrimiento de la actividad SOTCasa de la enzima *Bf*ArgF´ unido al hecho de que la ausencia del gen *argF´* en *B. fragilis* transforma este organismo en auxotrofo para la arginina indica que tanto succinilornitina y succinilcitrulina son intermediarios en la ruta de biosíntesis de la arginina en *Bacteroidetes*.

Lo expuesto en el anterior párrafo implica que la parte inicial de la ruta de biosíntesis de la arginina en los organismos que poseen actividad SOTCasa transcurre a través de intermediarios succinilados y no acetilados como ocurre en el resto de los organismos descritos hasta el momento.

Esta hipótesis es el punto de partida para los resultados expuestos en el siguiente capítulo.

Capítulo 4: "The use of N-succinyl-Lglutamate as substrate by ArgB of *Bacteroides fragilis* supports the existence in this organism of an arginine biosynthetic pathway using N-succinylated intermediates"

En este capitulo se presentan los datos obtenidos en el estudio de la actividad enzimática de los dos primeros genes supuestamente involucrados en al biosíntesis de la arginina en *B. fragilis* (*argX* and *argB*, Figura 37 de la tesis).

Ambos genes fueron clonados y las respectivas proteínas expresadas y purificadas, desafortunadamente la actividad enzimática de la proteína que hipotéticamente catalizaría el primer paso de la ruta, la succinilización del glutamato, no pudo ser cuantificada dado que la velocidad de la reacción espontánea era similar o superior a la de la reacción en presencia de la proteína codificada por el gen *Bf*argX (Figura 13). Posiblemente la enzima necesite unas condiciones diferentes a las utilizadas para ser activa, como por ejemplo algún cofactor desconocido de momento, o la interacción con otras proteínas.

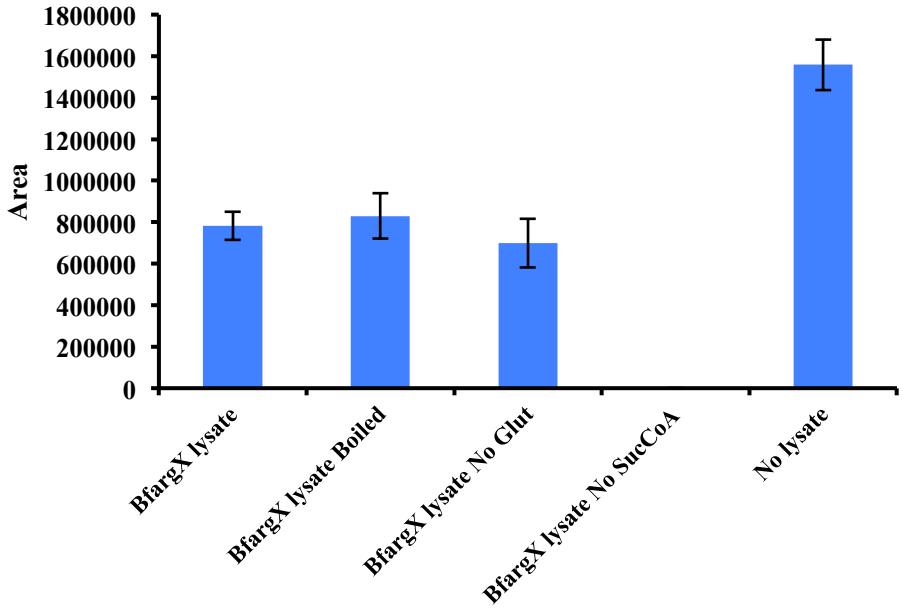


Figura 13: Ensayos iniciales para la determinación de la actividad enzimática de la proteína codificada por el gen *argX* de *B. fragils*.

Por consiguiente los esfuerzos se concentraron en el ensayo de la posible actividad succinilglutamato quinasa de la proteína codificada por el gen *argB* de *B. fragilis*. Para poder ensayar esta actividad, el posible sustrato, succinilglutamato, fue sintetizado químicamente (apartado 4.4 de la tesis) y dada el alto porcentaje de conversión (Figura 41 de la tesis), usado sin posterior purificación en los ensayos enzimáticos.

Como se muestra en la Figura 14, sólo el succinilglutamato puede ser utilizado como sustrato por la enzima *Bf*ArgB, lo cual demuestra que la biosíntesis de la arginina en *B. fragilis* transcurre a través de intermediarios succinilados y no

acetilados tal y como se había hipotetizado tras el descubrimiento de la actividad SOTCasa de *Bf*ArgF´ (apartado 4.5 de la tesis).

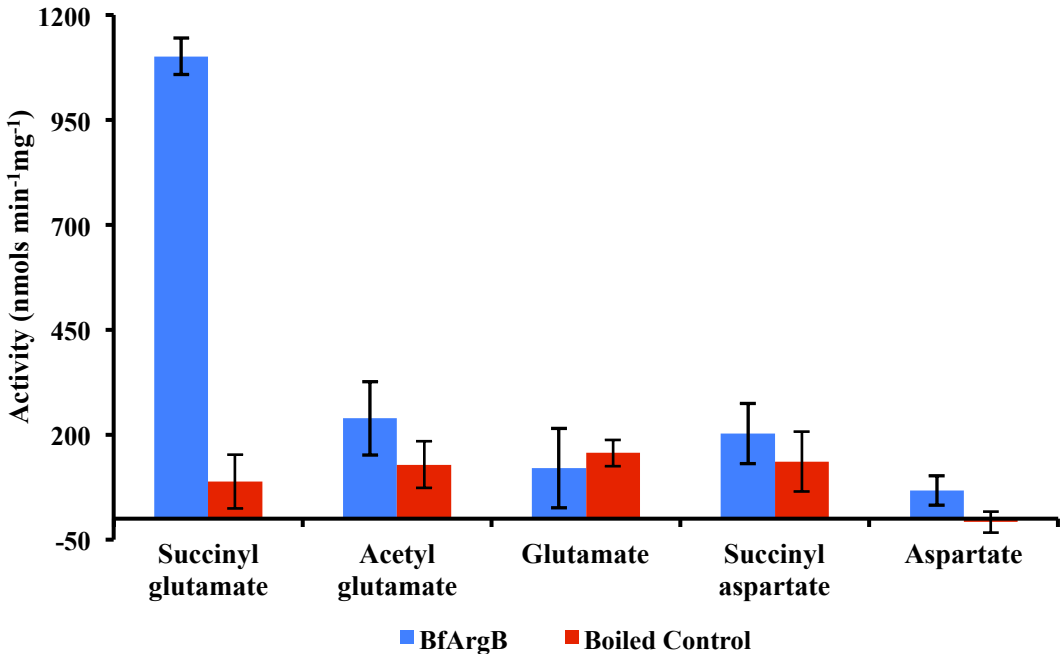


Figura 14: Ensayo de la actividad N-succinilglutamato quinasa de la enzima *Bf*ArgB'.

Usando el succinilglutamato sintetizado químicamente en el laboratorio, la actividad NSGK de *Bf*ArgB ha sido caracterizada y se han calculado los parámetros cinéticos (Apartado 4.6 de la tesis). Los parámetros cinéticos calculados para ambos sustratos de la reacción catalizada por BfArgB' se resumen en la siguiente tabla (Tabla 3):

Tabla 3: Resumen de los parámetros cinéticos calculados para ambos sustratos de la reacción catalizada por *Bf*ArgB'.

	Km (mM)	Vmax (µmols min^{-1} mg^{-1})
NSG	7.3 ± 0.44	52.1 ± 1.4
ATP	4.6 ± 0.25	54.7 ± 1.1

Dado que la actividad de la primera enzima involucrada en la ruta de biosíntesis *de novo* de la arginina en *B. fragilis* no pudo ser ensayada, y por lo tanto el posible efecto controlador que la arginina pudiera ejercer sobre ella no se pudo determinar, se estudió la posibilidad de que *Bf*ArgB' fuese el punto de control por retroalimentación negativa por la arginina en *B. fragilis*.

Para ello se seleccionaron las siguientes condiciones para el ensayo basándonos en los parámetros cinéticos estimados para la *Bf*ArgB' presentados en la anterior Tabla 3:

- Condiciones de saturación: 40mM NSG y 40mM ATP
- Condiciones cercanas a la Km: 10mM NSG y 6mM ATP
- Condiciones de no saturación: 3.6mM NSG y 2.3mM ATP

En estas tres condiciones se estudiaron los efectos de varias concentraciones de arginina (entre 0 y 50mM). Los resultados obtenidos se muestran a continuación (Figura 15).

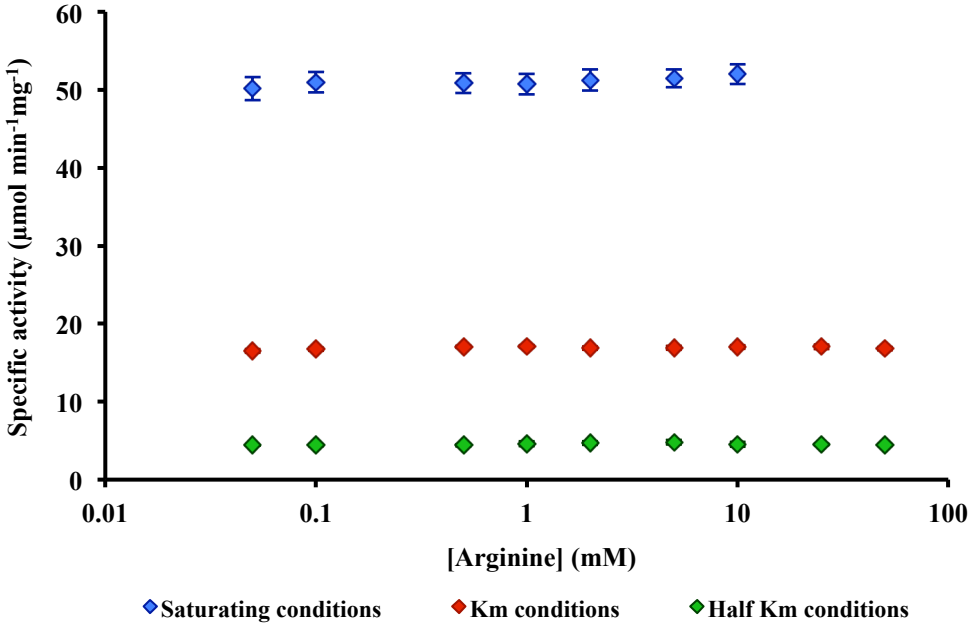


Figura 15: Estudios de inhibición por arginina de la actividad enzimática de *Bf*ArgB'.

Los resultados anteriormente expuestos, de los estudios de las actividades enzimáticas de las proteínas codificadas por los genes *argB'* y *argF"* nos permiten proponer la siguiente ruta metabólica para la biosíntesis *de novo* de la arginina en aquellos organismos dotados de SOTCasa en lugar de OTCasa o AOTCasa (Figura 16):

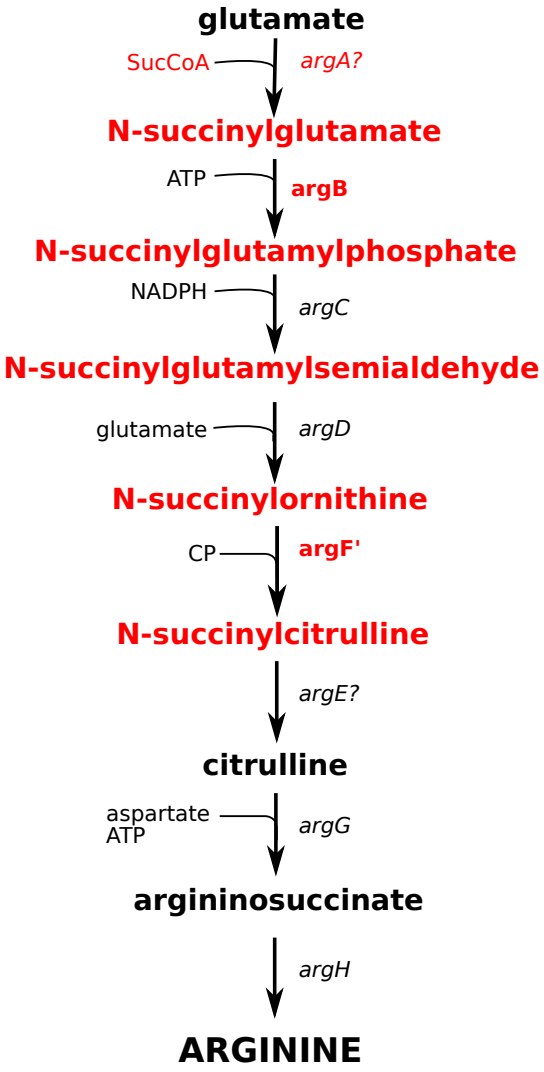


Figura 16:Representación esquemática de la ruta metabólica de biosíntesis de arginina en organismos dotados con SOTCasa.

Los resultados aquí resumidos se discuten en el Capítulo 5 de la tesis: **"General discussion: Novel pathways of arginine biosynthesis."**

A continuación se exponen las conclusiones obtenidas en el presente proyecto de tesis doctoral:

Conclusiones:

1. La enzima AOTCasa de X. campestris pv campestris, caracterizada en esta tesis, cataliza la formación de un nuevo intermediario en la ruta de biosíntesis de la arginina en éste organismo, demostrando la existencia de una nueva ruta biosíntetica para la arginina en los organismos dotados de AOTCasa y no OTCasa.

2. El gen *argF'* de *B. fragilis* codifica una enzima con actividad N-succinilornitina transcarbamilasa, otro nuevo miembro de la familia de las transcarbamilasas descubierto y descrito en este trabajo.

3. El gen *argB* de *B. fragilis* codifica una nueva enzima en la ruta de biosíntesis de la arginina en éste microorganismo: N-succinilglutamato quinasa.

4. La ruta de biosíntesis de la arginina en los microorganismos pertenecientes a la familia Bacteroidetes utiliza intermediarios succinilados en lugar de acetilados en la parte inicial de la ruta.

5. La identificación y caracterización bioquímica de las transcarbamilasas de *X. campestris* y *B. fragilis* permite la identificación de otros genes similares, que no codifican para OTCasas, mediante el uso de métodos bioinformáticos.