Volume 12 Number 15 1984 Nucleic Acids Research

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*

F.Van Vliet[1], R.Cunin[1], A.Jacobs[1], J.Piette[1], D.Gigot[3], M.Lauwereys[2], A.Piérard[3,4] and N.Glansdorff[1,4]

[1]Erfelijkheidsleer en Microbiologie, [2]Chemie der Proteinen, Vrije Universiteit Brussel, [3]Microbiologie, Université Libre de Bruxelles, and [4]Research Institute of the CERIA-COOVI, 1, ave. E. Gryson, B-1070 Brussels, Belgium

Received 17 May 1984; Revised and Accepted 6 July 1984

ABSTRACT

The complete nucleotide sequence of argF is presented, together with that of an operator-constitutive mutant. ArgF is compared with the other gene coding for ornithine carbamoyltransferase (OTCase) in E. coli K-12, argI, and with pyrB, encoding the catalytic monomer of aspartate carbamoyltransferase (ATCase). ArgF and argI appear very closely related having emerged from a relatively recent ancestor gene. The relationship between OTCase and ATCase appears more distant. Nevertheless, the homology observed between the two proteins (mainly in the polar domain) suggests a common origin.

INTRODUCTION

The enzymes ornithine carbamoyltransferase (OTCase, E.C. 2.1.3.3) and aspartate carbamoyltransferase (ATCase, E.C. 2.1.3.2) have been most intensively studied in E. coli. Both consist of trimeric complexes of catalytic subunits (1,2) which, in the case of ATCase, are assembled with regulatory polypeptides in a larger heteropolymer. Sequence homology has been uncovered between the aminoterminal domains of the catalytic monomers of ATCase and OTCase (3). As several bacteria display one OTCase gene (4,5) close to the pyrBI operon (encoding the catalytic and regulatory monomers of ATCase respectively - 6,7) it is possible that an ancestral carbamoyltransferase gene has been duplicated in near-tandem copies which have diverged in the course of evolution but have remained genetically linked at least in some organisms.

The pyrB-linked OTCase gene has been named argI in E. coli (4) ; curiously E. coli K-12 contains a second OTCase gene, argF (4), the product of which assembles with the argI monomer to form a family of trimeric isoenzymes (2). According to the expectations raised by this observation, heteroduplex analysis (8) and comparisons based on parts of the cognate amino acid (3,9) and nucleotide sequences (10) indicated that argF and argI should be highly homologous.

Table 1. Bacterial and bacteriophage strains used.

Bacteria	Origin
3000X111 thi Δ(pro lac argF)	F. Jacob
3000X111 thi thr argI Δ(pro lac argF)	this laboratory
3000X111 thi carB8 argI Δ(pro lac argF)	this laboratory
C600 Δ(pro lac argF) argI r^-,m^-	ref. 16
JM103 Δ(pro lac argF) supE, thi, strA, endA, sbcB15, hsdR4, F', traD36, $proAB^+$, $lacI^9ZM13$	ref. 17
Bacteriophages	
λpargF5 from lysogen N5	ref.18
M13 mp8	ref.19
M13 mp11	ref.20
Plasmid pMC20	ref.16

Here we present the complete sequence of argF and compare it with those of pyrB (11) and argI (12) which have become available since our initial comparative studies of their 5' regions. In addition, a mutation making the expression of argF constitutive is shown to alter the genetic element previously assumed to act as operator of that gene (10).

MATERIALS AND METHODS

Enzymes and Nucleoside Triphosphates

Restriction enzymes and T4 DNA ligase were purchased from Boehringer. [α or γ^{32}P]ATP was obtained from Amersham. T4 polynucleotide kinase and dideoxy nucleoside triphosphates were purchased from PL-Pharmacia.

Strains

See Table 1.

Fragment Mapping and DNA Sequencing

Fragment purification, mapping and labeling were as already described in ref.10. DNA sequences were determined either by the chemical degradation method (13) or by the enzymatic dideoxynucleotide terminating method (14).

Enzyme Assays

OTCase (E.C.2.1.3.3) was assayed as previously described (15).

RESULTS

Nucleotide Sequence of argF

The sequencing strategy is summarized in fig.1. The DNA source was plasmid pMC20, harbouring argF on a 1.8 kb BamHI insert. Segments close to the 5' region of argF were sequenced by chemical degradation (13). A 1160 b.p. HindIII-PstI fragment covering the rest of the gene and a 3' downstream sequence was subcloned into bacteriophage M13 mp8. Pure HindIII-PstI fragment was also restricted by TaqI and PvuII and the resulting subfragment respectively subcloned into M13 mp8 opened at its unique AccI site and into M13 mp11 cut with SmaI and PstI (for the PvuII-PstI subfragment) or with SmaI and HindIII (for the PvuII-HindIII subfragment). The M13 clones were used for sequencing by the dideoxynucleotide terminating technique (14).

The sequence of argF is displayed on fig.2. The results are in agreement with the amino acid sequence of several peptides coming from different parts of the enzyme (see legend of fig.2).

Comparison of argF with argI and pyrB

Fig.2 compares the nucleotide sequence of argF (top line) with that of argI (line 2). Homologies between the cognate amino acid sequences (lines 3 and 4) and that of the pyrB gene product (line 5) are indicated. Only in the latter case does the alignment of homologous residues require insertions or deletions which can be appreciated by the numbering of ATCase residues (310 in total against 333 for F and I OTCases). The regions of main homology

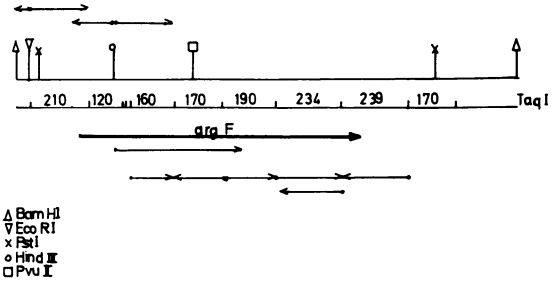
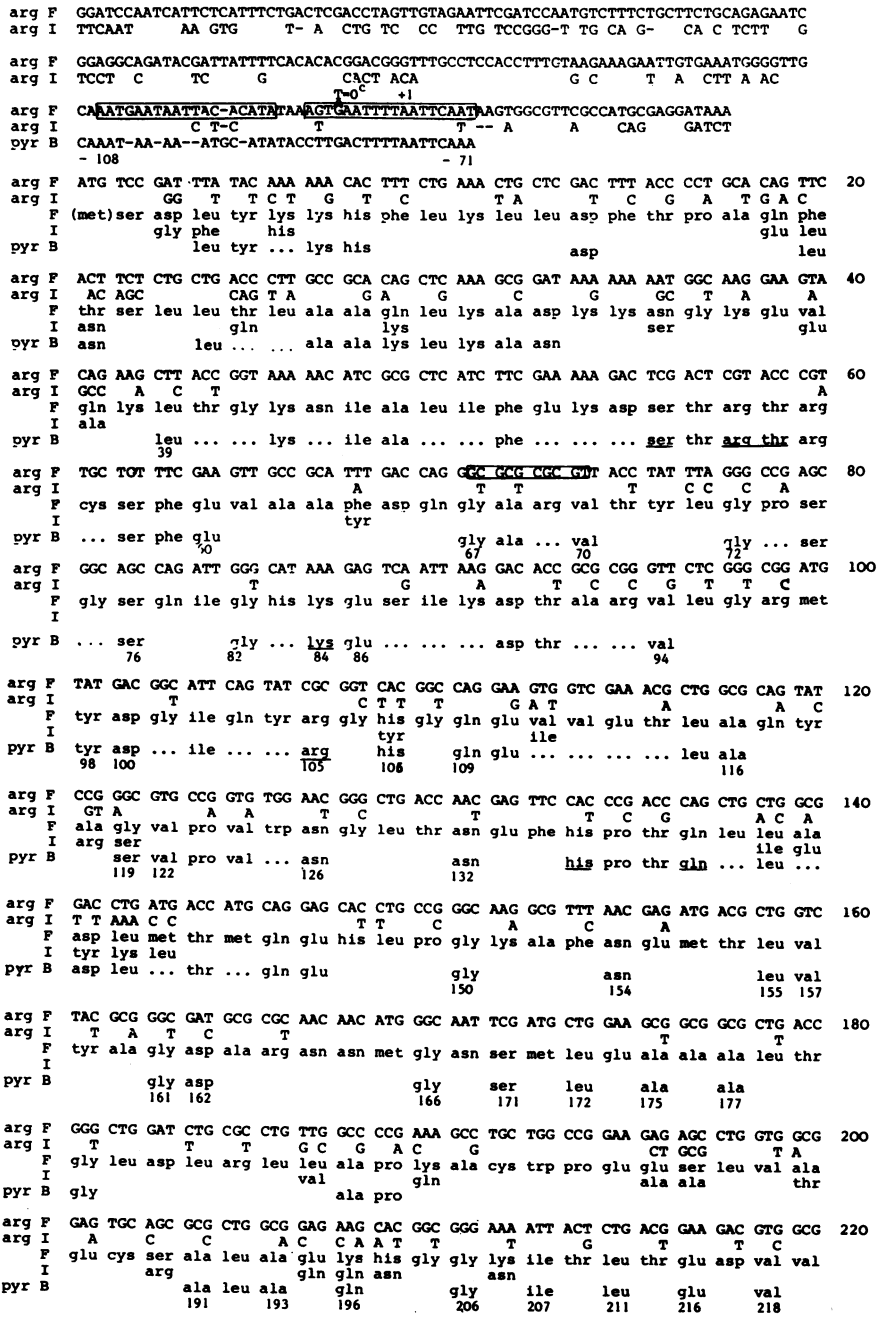


Figure 1. Restriction map and sequencing strategy of the argF gene.
The restriction map concerns the 1.8 kb BamHI insert of plasmid pMC20 (16) which contains the whole argF gene (solid arrow). The control and the N-terminal regions have been sequenced by the method of Maxam and Gilbert as described previously (10, see arrows above the restriction map ; 5' labeled ends are indicated by dots). The rest of the gene was sequenced by the chain terminating method (14) into the M13 mp8 vector (19). Arrows below the restriction map show the position and orientation of the sequenced inserts.



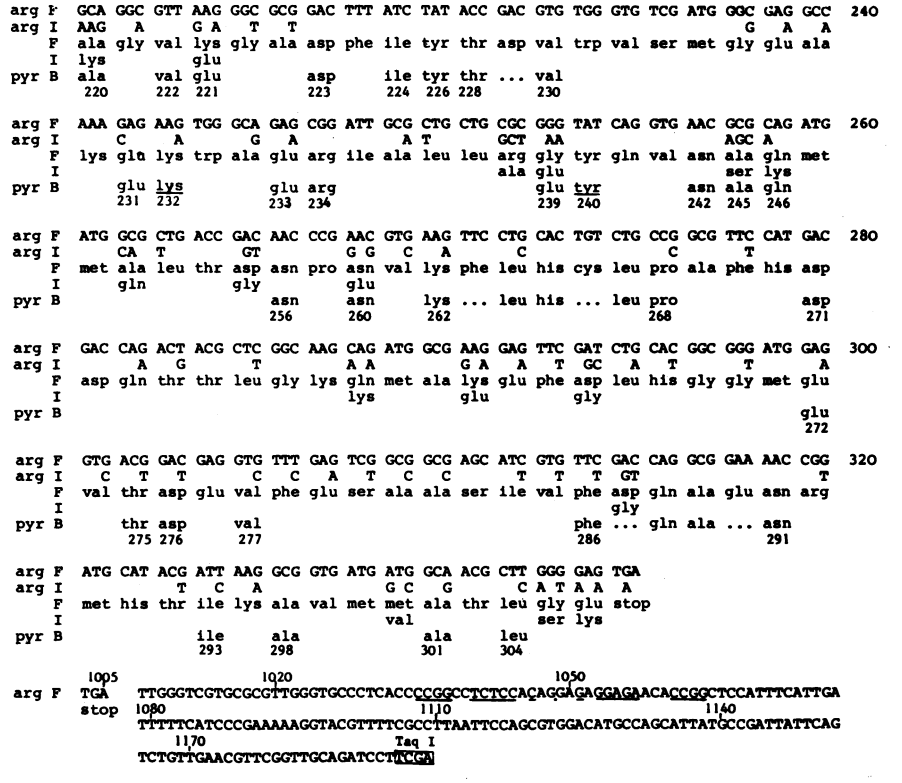


Figure 2. Nucleotide and corresponding amino acid sequence of argF - Comparison with argI and pyrB.

Only the differences between argF and argI are indicated. The numbering at the end of each line refers to the position of the last amino acid in the primary sequence. Amino acids of the pyrB ATCase subunit which are homologous with OTCase F/I are mentioned in the last line. Numbering of the ATCase-pyrB chain starts with ala-1 rather than met-1 to allow optimal alignment of the protein sequences. When pyrB residues are shifted with respect to argF and I sequences, the number of the first and last residues are indicated. The peptides analyzed correspond to the following residues: 2 to 34, 74 to 81, 100 to 108, 110 to 125, 153 to 162, 193 to 201, 225 to 232, 240 to 251, 253 to 256, 258 and 259, 290 to 297 and 330 to 334. Homologies and reverse homologies (in the 3' downstream region) were analyzed by computer using the CINTHOM and RINTHOM programs respectively (46). Underlined residues in the sequence of ATCase are those known to be involved in catalysis. The position of the operator-constitutive mutation of argF in the most distal of the tandem operators (boxes) is indicated.

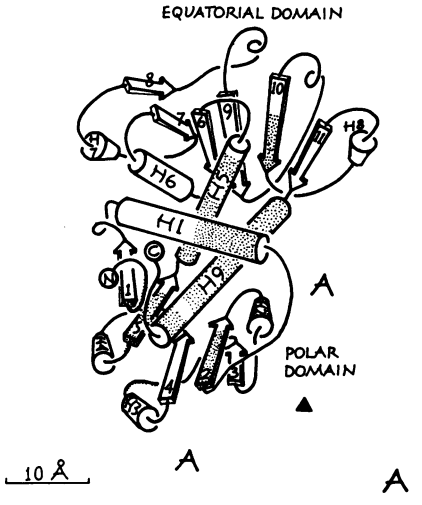


Figure 3. Three dimensional structure of the catalytic subunit of ATCase (encoded by the pyrB gene - ref. 21).
Regions of homology with OTCases F and I are dotted. Most of them are concentrated in the polar domain of the polypeptide chain and in the helixes 5 and 9 which are responsible for the interconnection between polar and equatorial domains.

between ATCase and OTCase are graphically displayed on fig.3 on the X-ray diffraction-based model (21) for the tertiary structure of the catalytic monomer of ATCase.

Structural predictions regarding the tertiary structure of OTCase have been worked out by the method of Chou and Fasman (22). They are compared on fig.4 with the actual structure of ATCase.

Codon Usage in argF

There are about 20,000 OTCase molecules per cell in the derepressed state, the respective contributions of argF and argI monomers being 0.2 and 0.5 % of total cell protein (6, and V. Stalon, pers. comm.). ArgF is thus a "moderately to highly" expressed gene (23). Codons which correspond to poorly represented tRNA's and may possibly limit the rate of translation (23,24,25) are however present in argF (Table 2) : 4 out of 11 arginine codons are CGG and 9 glycine codons out of 26 are GGG. Only two possible modulating arg codons (CGG and CGA) and 4 such glycine codons (one GGG and three GGA) are encountered in argI.

Analysis of the 3' Downstream Sequence

The 178 nucleotides following the translation termination codon of argF

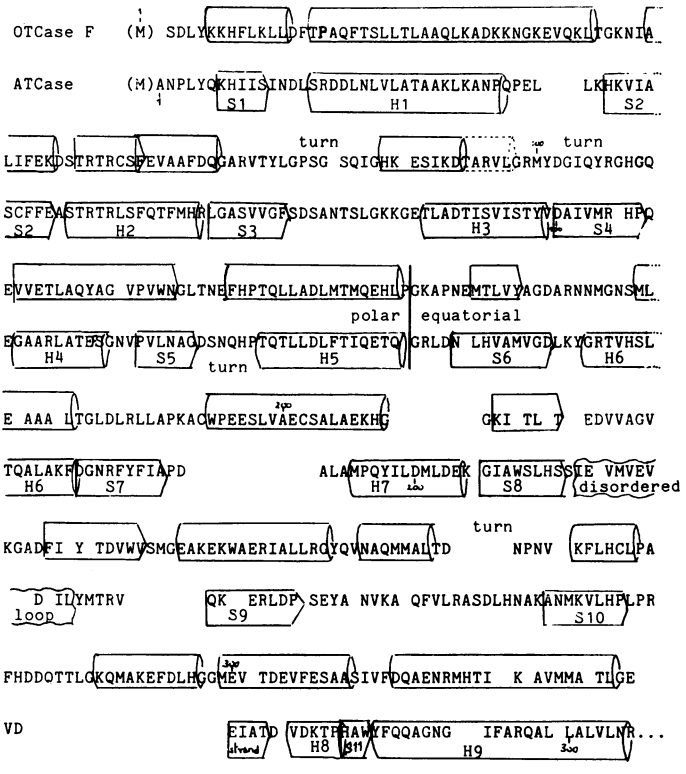


Figure 4. Prediction of the secundary structure of OTCase F.
The Chou and Fasman method (22) has been used with the help of a computer to predict secundary structures (α helices, β sheets and turns) within OTCase F (line 1). It is compared to the actual secundary structure of ATCase (line 2). Cylinders represent α helices and arrows β sheets.

have been screened for potential stem-loop structures that could be associated with transcription termination. Only one potentially stable (-14.8 Kcal) structure could be found, involving the nucleotides between positions 1035 and 1064.

Isolation and Characterization of an argF Operator-Constitutive Mutation

Mutants with high OTCase specific activity can be selected from strains defective in carbamoylphosphate synthesis (15) ; these mutants rely on phosphorolysis of citrulline for carbamoylphosphate production. The most frequent types are constitutive argF and argI mutants or argG leaky strains (15). We isolated argF constitutive mutants from λC_I857 pargF5 lysogen of strain 3000X111 carB8 Δ(pro lac argF) ; out of 25 citrulline utilizing isolates,

Table 2. Codon usage in argF.

NUMBER OF IDENTIFIED CODONS= 334
NUMBER OF UNIDENTIFIED CODONS= 0
CALCULATED MOLECULAR WEIGHT= 36850.39

TTT	PHE	6	1.8%	TCT	SER	2	.6%	TAT	TYR	6	1.8%	TGT	CYS	1	.3%
TTC	PHE	8	2.4%	TCC	SER	1	.3%	TAC	TYR	2	.6%	TGC	CYS	3	.9%
TTA	LEU	2	.6%	TCA	SER	1	.3%	TAA	---	0	---	TGA	---	1	---
TTG	LEU	1	.3%	TCG	SER	4	1.2%	TAG	---	0	---	TGG	TRP	4	1.2%
CTT	LEU	3	.9%	CCT	PRO	1	.3%	CAT	HIS	3	.9%	CGT	ARG	2	.6%
CTC	LEU	5	1.5%	CCC	PRO	0	0.0%	CAC	HIS	7	2.1%	CGC	ARG	5	1.5%
CTA	LEU	0	0.0%	CCA	PRO	0	0.0%	CAA	GLN	0	0.0%	CGA	ARG	0	0.0%
CTG	LEU	25	7.5%	CCG	PRO	9	2.7%	CAG	GLN	15	4.5%	CGG	ARG	4	1.2%
ATT	ILE	6	1.8%	ACT	THR	4	1.2%	AAT	ASN	2	.6%	AGT	SER	0	0.0%
ATC	ILE	4	1.2%	ACC	THR	12	3.6%	AAC	ASN	10	3.0%	AGC	SER	5	1.5%
ATA	ILE	0	0.0%	ACA	THR	0	0.0%	AAA	LYS	12	3.6%	AGA	ARG	0	0.0%
ATG	MET	15	4.5%	ACG	THR	7	2.1%	AAG	LYS	11	3.3%	AGG	ARG	0	0.0%
GTT	VAL	4	1.2%	GCT	ALA	0	0.0%	GAT	ASP	5	1.5%	GGT	GLY	2	.6%
GTC	VAL	2	.6%	GCC	ALA	6	1.8%	GAC	ASP	14	4.2%	GGC	GLY	15	4.5%
GTA	VAL	1	.3%	GCA	ALA	6	1.8%	GAA	GLU	9	2.7%	GGA	GLY	0	0.0%
GTG	VAL	13	3.9%	GCG	ALA	25	7.5%	GAG	GLU	15	4.5%	GGG	GLY	9	2.7%

twenty exhibited a high (more than 200 µM/hr/mg protein) OTCase specific activity in the presence of arginine and uracil. Lysates were produced by thermoinactivation of the $\lambda C_I 857$ repressor and used to transduce C600 argF argI to arginine auxotrophy. Four transductants with high OTCase specific activity were analyzed further. When their nucleotide sequence was established only one of them (argF-O^c-3) was found to harbour a mutation in the 5' non-coding region of argF (Fig.2). The alterations responsible for the constitutivity of the other three mutants have not been identified although their respective sequences have been established as far as within the λint gene. They could be mutations creating a promoter far enough upstream from the operator to make initiation of transcription relatively insensitive to repression (26,27).

The novel joint created by recombination between λattP and a secundary chromosomal att site during the insertion of the prophage ancestor of λpargF5 was found immediately after a C residue (circled below) occupying position + 60, upstream of the argF transcription startpoint ; in the secundary site 15 residues (underlined) match the 21 bases of the core-consensus sequence derived from λattB and P (28) : CAATTCTTTCTTA©AAAGGTG (reading away from argF on the non-coding strand). Since a λpargF5 lysogen shows normal expression and regulation of the prophage-born argF (Table 3), participation to argF expression of a putative promoter centered on nucleotide -110 (29) and of an hyphenated sequence between residues -46 and -75 is ruled out, leaving as sole controlling elements the promoter and the two "ARG boxes" (argF01 and argF02) we previously described (10).

The mutation in argF-O^c-3, a spontaneous GC → TA transversion, results in a 12-fold drop of argF repressibility (Table 3).

Table 3. Enzymatic values of OTCase F in the constitutive mutant and its wild-type parent.

Strain	OTCase F specific activity (µM citrulline/hr/mg prot.)
3000X111 *argI argF* (λ*pargF5*) *argR*	225
" " (λ*pargF5*) *argR*$^+$	3.8
" " (λ*pargF5O*^{c}*-3*) *argR*	553
" " (λ*pargF5O*^{c}*-3*) *argR*$^+$	116

The *argR* strains were isolated on canavanine resistant derivatives (47) of their *argR*$^+$ parent.

DISCUSSION

The argF Gene

The present analysis completes the sequence of the *argF* gene and confirms the previous assignment of the operator sequence by locating a mutation within one of the two 18 b.p. long palindromes overlapping the *argF* promoter (Fig.2). We have found similar tandem operators, also separated by 3 b.p., in the *argECBH* operon (30), in the *carAB* operon (27,31) and in *argI* (10). The present mutation affects a GC base pair which is strongly conserved among the ARG boxes and is also involved in their dyad symmetry. This is in striking contrast with the GC ⟶ AT substitution, which, at the 14th position in box *argECBH-1* (30), alters a non-conserved base pair. Both mutations may be explained by the loss of an important contact between the DNA sequences and the repressor but the *argF O*^{c} mutation in particular emphasizes the importance of dyad symmetry in this interaction.

A putative terminator of *argF* transcription is indicated on Fig.2. This assignment is based on the possibility of forming a stem-loop secundary structure followed by a stretch of nucleotides (TCCATTTCATTGA) bearing extensive similarity with sequences found in two rho-dependent terminators : TAATTCTCATTAG in *trpt* and TATTCATTA in rrnBT$_2$ (see ref. 32). This structure combines a relatively stable stem-loop (-14.8 Kcal) - a feature of rho-independent-terminators - with sequences reminiscent of rho-dependent terminators. The actual role of this sequence is being investigated further.

ArgF Versus ArgI

The sequence now available allow us to compare *argF* and *argI*, from 190

nucleotides upstream from the transcription startpoint down to the translation stop codon. The present data confirm that the homology between argF and argI is very extensive indeed, in fact much more than had been assumed on the basis of RNA to DNA hybridization experiments (33) ; it is however not homogeneously distributed over the whole length of the gene. Indeed, while 78.1 % overall homology is observed at the nucleotide level and 86 % at the amino acid level, we note a stretch of almost perfect homology running from codon 45 to codon 70. This stretch is followed by a sequence displaying alternative purines and pyrimidines (boxed on Fig.2 ; note that a similar stretch is delineated by residues 165 and 178). In eukaryotes such sequences have been suggested to play a role in the mechanism of gene conversion, possibly by promoting and/or terminating recombination or strand exchanges between genes (34,35). Part of the observed homology between argF and argI might be due to such a conversion event. The occurrence of putative converted segments bound by pu-py stretches in organisms as different as pro- and eukaryotes would suggest that those sequences promote or terminate strand-exchange because of basic structural property : formation of Z-DNA may be envisaged as an interesting possibility.

If we except codons 45 to 70 and calculate the nucleotide and amino acid homologies between the two genes on the rest of the sequence, we find 77 and 85 %, respectively. These values may now be compared with the corresponding homologies observed between trp genes in the closely related species E. coli and Shigella dysenteriae and the slightly more distant species E. coli and Salmonella typhimurium. E. coli and Shigella display 95.6 % (nucleotides) and 97.9 % (amino acids) for trpG (36), while E. coli and Salmonella show 80.2 and 87.3 % for trpE (37), 82.5 and 95.4 % for trpG (36) 85 and 96 % for trpB (38) but only 75.2 and 85.1 % for trpA. These values reflect, in unknown proportions, the degree of conservation imposed on a particular protein by molecular constraints, such as dictated by interaction with other proteins. The spreading of these values show that, on such basis alone, it would not be possible to locate an ancestral divergence between argF and argI with great precision. It is possible that argF is the result of translocation of argI from a related species into E. coli as inferred (40,41) from the presence of two IS1 elements flanking argF. However the still high degree of argF-argI homology does not exclude that the divergence has occurred within a branch leading to present-day E. coli, with possible interstrain transfer occurring later (10,42).

ArgF and argI appear as "moderately to highly expressed genes" (23, and

Results). A correlation has been established between the relative concentrations of isoacceptor tRNA's and the use of the corresponding codons in the mRNA of highly expressed genes (24,25). Codons corresponding to poorly represented tRNA species might limit the rate of translation (ibid.,23,24) ; argF contains 13 such codons and argI 6 ones. It is unlikely that this difference is responsible for the fact that E. coli produces three times as much I than F monomers since amplification of argF on a multicopy plasmid (16) leads to a specific activity (6000 µM/hr/mg protein) corresponding to 5 % of total cell protein. This suggests that the rate of translation of argF is not substantially restricted by the use of minor tRNA's. We would therefore tend to conclude that codons corresponding to poorly represented tRNA species display only weak modulating effects - if any - in situations where one attemps to assess their direct physiological impact. It remains however possible that, in the long run, even weak effects may determine an evolutionary trend towards a correlation between the relative concentrations of tRNA's and cognate codons in highly expressed genes.

OTCase versus ATCase

The comparison between the argF-argI couple and pyrB concerns an event of a much more primeval nature than the origin of argI : the emergence of two basic metabolic functions respectively involved in amino acid and pyrimidines autotrophy. If this event consisted in enzyme differentiation by recruitment (44) for one of these pathways of an ancestral carbamoyltransferase displaying some substrate ambiguity, it would not be unexpected to find evidence of this in the spatial structure of these enzymes and in their primary sequence As a matter of fact both types of homologies have been detected.

The overall homology between argF or argI and pyrB is 35 to 40 %. The main regions of homology are unevenly distributed, occurring mostly in the polar moiety (which binds carbamoylphosphate) and in helical regions (H5 and H9) joining the polar and equatorial domains (Fig.3). There is no homology in the region which, in the catalytic monomer of ATCase, contains a secundary CTP binding site. Extensive backbone homology, mainly in the polar domain, is attested by the comparison (see Fig.4) of the structural predictions made for OTCase with the actual structure of ATCase. Besides, sequence homologies between the two proteins are found in similarly structured domains (see Figs 3 & 4).

Three dimensional analysis of OTCase by X-ray diffraction is required in order to further evaluate the evolutionary relationship between carbamoyltransferases but the available data suggest strongly that present-day

OTCases and ATCase arose by divergent evolution from an ancestral transferase.

In general, proteins contain relatively little arginine when considering that there are six codons for this amino acid. On this basis, Jukes (45) has suggested that arginine would have been a "late intruder" having replaced ornithine in protein synthesis in the course of evolution. As arginine is synthesized from ornithine, this theory would suggest that OTCase has been recruited from an ancestral carbamoyltransferase originally involved in aspartate carbamoylation but endowed with some substrate ambiguity.

ACKNOWLEDGEMENTS

This work has been supported by Belgian National Research Institutions (IRSIA-IWONL and FRFC-FKFO) as well as by an "Action Concertée" (Onderling Overlegde Aktie) between the Free Universities of Brussels and the Belgian Government.

We thank Dr. Marc Collet for help with the Chou and Fasman computer program.

REFERENCES

1. Rosenbusch, J.D. and Weber, K. (1971) J. Biol. Chem. 246, 1644-1657.
2. Legrain, C., Halleux, P., Stalon, V. and Glansdorff, N. (1972) Eur. J. Biochem. 27, 93-102.
3. Gigot, D., Glansdorff, N., Legrain, C., Piérard, A., Stalon, V., Konigsberg, W., Caplier, I., Strosberg, A.D. and Hervé, G. (1977) FEBS Lett. 81, 28-32.
4. Glansdorff, N., Sand, G. & Verhoef, C. (1967) Mut. Res. 4, 743-751.
5. Jacoby, G.A. (1971) J. Bacteriol. 108, 645-651.
6. Feller, A., Piérard, A., Glansdorff, N., Charlier, D. and Crabeel, M. (1981) Nature 292, 370-373.
7. Wild, J.R., Foltermann, K.F., Roof, W.D. and O'Donovan, G.A. (1981) Nature 292, 373-375.
8. Kikuchi, A. and Gorini, L. (1975) Nature 256, 621-624.
9. Piette, J., Gigot, D., Lauwereys, M., Cunin, R., Glansdorff, N., and Piérard, A. (1982) Arch. Intern. Physiol. Biochim. 90, B211-B212.
10. Piette, J., Cunin, R., Van Vliet, F., Charlier, D., Crabeel, M., Ota, Y., and Glansdorff, N. (1982) The EMBO J. 1, 853-857.
11. Hoover, T.A., Roof, W.D., Foltermann, K.F., O'Donovan, G.A., Bencini, D.A. and Wild, J.R. (1983) Proc. Natl. Acad. Sci. USA 80, 2462-2466.
12. Bencini, D.A., Houghton, J.E., Hoover, T.A., Foltermann, K.F., Wild, J.R. and O'Donovan, G.A. (1983) Nucl. Ac. Res. 11, 8509-8518.
13. Maxam, A. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
14. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
15. Legrain, C., Stalon, V., Glansdorff, N., Gigot, D., Piérard, A. and Crabeel, M. (1976) J. Bacteriol. 128, 39-48.
16. Crabeel, M., Charlier, D., Cunin, R. and Glansdorff, N. (1979) Gene 5, 207-231.

17. Messing, J., Crea, R. and Seeburg, P.H. (1981) Nucl. Ac. Res. 9, 309-321.
18. Legrain, C., Stalon, V. and Glansdorff, N. (1976) J. Bacteriol. 128, 35-38.
19. Hu, N.T. and Messing, J. (1982) Gene 17, 271-277.
20. Messing, J. and Vieira, J. (1982) Gene 19, 269-275.
21. Monaco, H.L., Crawford, J.L. and Lipscomb, W.N. (1975) Proc. Natl. Acad. Sci. USA 75, 5276-5280.
22. Chou, P.Y. and Fasman, G.D. (1974) Biochemistry 13, 211-222 and 222-245.
23. Grosjean, H., and Fiers, W. (1982) Gene 18, 199-209.
24. Ikemura, T. (1981a) J. Mol. Biol. 146, 1-21.
25. Ikemura, T. (1981b) J. Mol. Biol. 151, 389-409.
26. Reznikoff, W., Miller, J.H., Scaife, J.G. and Beckwith, J.R. (1969) J. Mol. Biol. 43, 201-213.
27. Cunin, R., Eckhardt, T., Piette, J., Boyen, A., Piérard, A. and Glansdorff, N. (1983) Nucl. Ac. Res. 11, 5007-5019.
28. Ross, W. and Landy, A. (1983) Cell 33, 261-272.
29. Moore, S., James, E., Fareed, G. and James, P. (1978) Gene 4, 261-278.
30. Piette, J., Cunin, R., Boyen, A., Charlier, D., Crabeel, M., Van Vliet, F., Glansdorff, N., Squires, C. and Squires, C.L. (1982) Nucl. Ac. Res. 10, 8031-8047.
31. Piette, J., Nyunoya, H., Lusty, C.J., Cunin, R., Weyens, G., Crabeel,M., Charlier, D., Glansdorff, N. and Piérard, A. (1984) Proc. Natl. Acad. Sci. USA, in the press.
32. Kingston, R.E. and Chamberlin, M.J. (1981) Cell 27, 523-531.
33. Sens, D., Natter, W. and James, E. (1977) Cell 10, 275-285.
34. Slightom, J.L., Blechl, A.E. and Smithies, O. (1980) Cell 21, 627-638.
35. Nordheim, A. and Rich, A. (1983) Proc. Natl. Acad. Sci. USA 80, 1821.
36. Nichols, B.P., Miozzari, G.F., van Cleemput, M., Bennet, G.N. and Yanofsky, C. (1980) J. Mol. Biol. 142, 503-517.
37. Yanofsky, C. and van Cleemput, M. (1982) J. Mol. Biol. 155, 235-246.
38. Crawford, I.P., Nichols, B.P. and Yanofsky, C. (1980) J. Mol. Biol. 142, 489-502.
39. Nichols, B.P. and Yanofsky, C. (1979) Proc. Natl. Acad. Sci. USA, 76, 5244-5248.
40. Hu, M. and Deonier, R.C. (1981) Mol. Gen. Genet. 181, 222-229.
41. York, M.K. and Stodolsky, H. (1981) Mol. Gen. Genet. 181, 230-240.
42. Riley, M. and Anilionis, A. (1978) Ann. Rev. Microbiol. 32, 519-560.
43. Gouy, M. and Gautier, C. (1982) Nucl. Ac. Res. 10, 7055-7074.
44. Jensen, R.A. (1976) Ann. Rev. Microbiol. 30, 409-425.
45. Jukes, T.M. (1973) Biochem. Biophys. Res. Comm. 53, 709-714.
46. Pustell, J. and Kafatos, F.C. (1982) Nucl. Ac. Res. 10, 51-59.
47. Novick, R.P. and Maas, W.K. (1961) J. Bacteriol. 81, 236-240.