The DNA sequence of *argI* from *Escherichia coli* K12

David A.Bencini⁺, John E.Houghton, Timothy A.Hoover, Karen F.Foltermann, James R.Wild* and Gerard A.O'Donovan

Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA

Received 24 June 1983; Revised and Accepted 28 October 1983

ABSTRACT

The *argI* gene from *E. coli* K12 has been sequenced. It contains an open reading frame of 1002 bases which encodes a polypeptide of 334 amino acids. Three such polypeptides are required to form the functional catalytic trimer (c_3) of ornithine transcarbamoylase (OTCase-1, EC 2.1.3.3). The molecular mass of the mature trimer deduced from the amino acid sequence is 114,465 daltons. An altered form of *argI* was produced when a 1.6 kilobase *DdeI* fragment was subcloned into the *HincII* site of plasmid pUC8 extending the open reading frame an additional 20 nucleotides. It has been previously reported that the amino-terminal region of the respective polypeptides of *argI*, *argF*, and *pyrB* of *E. coli* possessed significant homology. In contrast, the homologous promoter/operator regions of *argI* and *argF* did not appear to share any homologies with *pyrB*. However, a closer scrutiny of the nucleotide sequence immediately preceding the *pyrBI* attenuator revealed a remarkable similarity to the *argI* and *argF* control region.

INTRODUCTION

The enzyme ornithine transcarbamoylase (OTCase; carbamoyl phosphate: L-ornithine carbamoyltransferase, EC 2.1.3.3) catalyzes the formation of citrulline and phosphate from ornithine and carbamoyl phosphate in the sixth reaction unique to the arginine biosynthetic pathway. There are two genes, *argI* and *argF*, which encode OTCase in *Escherichia coli* K12 (1). Since OTCase is a trimer, these two genes result in the synthesis of four isoenzymes (III, IIF, IFF, and FFF) (2). The *argI* gene is located at 96.5 min on the *E. coli* chromosome and is situated about 4000 base pairs from the *pyrBI* operon which encodes aspartate transcarbamoylase (ATCase; carbamoyl phosphate: L-aspartate carbamoyltransferase, EC 2.1.3.2) (3). The *argF* gene is situated at 6 min and is not closely linked to the *argI*-*pyrBI* cluster (4). Several other isofunctional enzymes have been reported in *E. coli* but each family is characterized by different regulatory properties (e.g. aspartokinase, homoserine dehydrogenase, DAHP synthetase, and chorismate

mutase) (5). The only other known duplications of protein function in *E. coli* involve the OTCases and the protein chain elongation factors encoded by tufA and tufB (6).

The carbamoyl transferases encoded by argI, argF, and pyrB are structurally similar. The OTCases encoded by argI and argF are trimeric enzymes of similar size (about 110,000 daltons) but differ considerably in physical characteristics such as surface charge and heat stability (2). The ATCase holoenzyme consists of 6 pyrB-encoded (catalytic, c) polypeptides and 6 pyrI-encoded (regulatory, r) polypeptides in a 2(c₃):3(r₂) configuration (7). ATCase is still functional as a catalytic trimer (c₃) when the regulatory polypeptide is removed by biochemical (7) or genetic means (3). Similar to the OTCase isoenzymes, the catalytic trimer exhibits Michaelian kinetics and has a molecular weight of approximately 100,000 daltons (2). These genetic, structural, and functional similarities support the hypothesis that these enzymes have diverged from a common ancestral carbamoyl transferase (8).

An evaluation of the evolution of *E. coli* carbamoyl transferases and their genes is aided by detailed studies of the crystallographic structure (9,10), biochemical modification (11), DNA sequence (12), and corroborating amino acid sequence (13) of aspartate transcarbamoylase. Detailed comparative studies, however, require information concerning the primary structure of OTCase or the DNA sequence of argI or argF. Piette et al. (14) and Moore (15) have previously established the DNA sequence of the promoter regions of argI and argF and as much as 20% of the structural portion of the argI gene. In this study we have established the DNA sequence of the entire structural portion of the argI gene and deduced the amino acid sequence of the OTCase-1 encoded by argI (distinguished from the argF gene product or isozymic hybrids.).

MATERIALS AND METHODS

Cloning Procedure

Plasmid pPBh102 (3) contains both argI and pyrBI and was used to construct smaller subclones for use in DNA sequencing. Various plasmid inserts are depicted in Figure 1. All plasmids except pAI201 were derivatives of pBR322. Fragments which possessed the blunt-ended HpaI site were inserted into the PvuII site of pBR322. The smallest subclone, pAI201, contains a 1.6 bp DdeI fragment inserted into the polylinker region of pUC8 (16). The fragment was made blunt-ended by a

DNA polymerase fill reaction using the Klenow fragment of DNA Polymerase I (17) then ligated into the HincII site of the polylinker within the lacZ' gene (Figure 2). The plasmid was then selected by insertional inactivation of the lacZ' gene using host JM 83 on media containing ampicillin and X-gal as described by Vierira and Messing (16). All of these plasmids were capable of complementing an Arg⁻ background, TB2 [E. coli K12: Δ (argI-pyrBI) argF⁻] (3).

DNA Sequencing

Plasmid pAI201 was used to sequence argI except for the final eight nucleotides (see Figures 1 and 3). Both strands have been sequenced except for the initial 64 nucleotides which have been sequenced independently in two other laboratories (14,15) and the final eight nucleotides of the sequence which have been sequenced only on the sense strand. The DNA was sequenced by modification of the method of Maxam and Gilbert (18). These modifications will be described in detail elsewhere (manuscript in preparation) but highlights of the method are as follows. The DNA fragments were labeled at the 5' end with polynucleotide kinase or at the 3' end using the Klenow fragment of DNA polymerase I (17). The fragments were isolated on a thiol-crosslinked 5% polyacrylamide gel (19), excised, and dissolved by adding β -mercaptoethanol to 20% of the total volume. The fragments, in a 800-1200 μ l volume, were extracted into several 200 μ l volumes of butanol containing 1% cetyltrimethylammonium bromide. After eliminating the aqueous phase by brief centrifugation, the DNA (in the organic phase) was repartitioned into 200 μ l of 1 M NaCl. The upper (organic) phase was removed and the DNA precipitated by adding 2.5 volumes of ethanol. The pelleted DNA was then washed twice with 70% ethanol at room temperature, dried in vacuo, and resuspended in 50 μ l distilled and deionized water. The C+T, G+A and A>C modification reactions were used to obtain the DNA sequence from fragments uniquely labeled at the 3' and 5' end. The C+T reaction was used essentially as described previously (18). The G+A reaction was adjusted to 0.1 M NaOH and placed on ice until used in the strand scission reaction. The A>C reaction was performed in a 20 μ l volume. Acetic acid was added to adjust the reaction to 0.1 N base then placed on ice until used in the strand scission reaction. Strand scission required the addition of 100 μ l 1 M piperidine followed by incubation at 90°C for 30 min. To avoid lengthy and tedious multiple lyophilizations, the samples were precipitated by

the addition of 1.2 ml butanol, followed by a 1 min spin in a microcentrifuge (10000 x g). The butanol was removed, the DNA pellet was rinsed with 1 ml 70% ethanol, and the DNA was dried for 5 min in vacuo before being applied to sequencing gels as described by Smith and Calvo (20).

RESULTS AND DISCUSSION

The argI gene is located approximately 4000 base pairs from pyrBI (3,12; see Figure 1) at 96.5 minutes on the E. coli map. Plasmid pAI201, containing the argI⁺ 1.6 kilobase (kb) DdeI fragment from pAI104, was able to complement an argI⁻ argF⁻ background (Figures 1 and 2). In contrast, plasmids containing restriction endonuclease fragments derived from digestion with PstI or BstEII failed to complement the same background. DNA sequence analysis revealed an open reading frame which corresponded not only to the published N-terminus of argI (14,15) but also to homologous regions of pyrB (3,12). The translational reading frame and promoter region was similar to that reported earlier (14,15).

The majority of the DNA sequence for argI was established from pAI201. This plasmid contained argI in an orientation opposite to that of the lacZ' gene (Figure 2). The open reading frame within pAI201 was found to extend from argI into the cloning vehicle. This reading frame contained an in-phase stop codon 20 nucleotides after the DdeI site which defined the boundary between the insert and the cloning vehicle (Figure 3). DNA sequence analysis using pAI104 indicated that the wild-type gene extended only five nucleotides beyond the DdeI restriction site before encountering a stop signal. Thus, the altered argI gene replaced the C-terminal dipeptide -ser-lys-COOH of the wild-type OTCase with the heptapeptide -arg-arg-ile-pro-gly-asn-ser-COOH (see Figure 3). The substitution of the C-terminal dipeptide suggests that the wild-type ser-lys-COOH sequence may not be essential for the the function of OTCase. As can be seen from Figure 3, the addition of this new sequence also resulted in a novel fusion of the argI gene with the polylinker of pUC8 and adds several new restriction sites to the sequence of the altered argI gene.

The open reading frame of the wild-type argI gene (Figure 4) was found to contain 1002 nucleotides (excluding the stop codon but including the AUG start codon) and 1017 nucleotides for the altered argI

Figure 1. The cloning and sequencing of *argI*. The figure illustrates (top) the chromosomal location of *argI*, *pyrBI*, and *argF*; (middle) a restriction map of *argI*-containing fragments with plasmid nomenclature; (bottom) the restriction map generated from plasmid pAI201 used in DNA sequencing. Straight arrows designate regions sequenced using a 5' label, and broken arrows designate regions sequenced using a 3' label.

gene. The mature polypeptide, not including the amino-terminal methionine (8), is 333 amino acids in length with a molecular mass of 38,155 daltons. The corresponding trimeric enzyme is 114,465 daltons. Legrain et al. estimated the size of OTCase-1 to be 110,000 daltons by SDS gel electrophoresis of the denatured form of the enzyme and 140,000

Figure 2. The construction of plasmid pAI201. A, *DdeI* digestion; B, DNA polymerase I fill; C, *HincII* digestion; and D, blunt-end ligation.

Figure 3. The altered C-terminal region of OTCase produced by subcloning *argI* into plasmid pUC8. A and B, *Dde*I digestion and DNA polymerase I fill of pAI201. The region shown encodes the C-terminal region of the OTCase polypeptide. C, *Hinc*II digestion of pUC8. This restriction site is part of the "polylinker" within the *lacZ'* structural gene. D, blunt-end ligation.

daltons by molecular sieve chromatography (2). The data presented here support their contention that the chromatographic estimation of molecular weight was high, possibly due to the fact that the geometry of OTCase is considerably different from the standard proteins used to calibrate the column.

The DNA sequence reported herein is identical to the partial DNA sequence reported for the structural portion of the *argI* gene by Piette et al. (14) and Moore (15). The latter study reported 207 nucleotides of the structural portion of the *argI* gene immediately following the promoter sequence. The amino acid sequence of the OTCase polypeptide deduced from the DNA sequence also agrees with the amino-terminal sequence of the first 37 residues reported by Gigot et al. (8) determined from Edman degradation of the amino-terminus of the polypeptide (except for his-5, ala-27, and lys-35 reported as arg, ser, and glx by amino acid sequencing). Recently, Piette et al. (21) has reported the deduced amino acid sequence of the first 86 residues of the *argI*-encoded polypeptide (although no DNA sequence data was given). The deduced amino acid sequence reported here is identical to that given by Piette et al. except that residues #39, 81, and 85 (glu, ser, and his in our sequence) are reported as gln, arg, and pro in that study.

a	-35	-10	+1	+25																					
	ATGCTTTAGA CTTGCAAAATG AATAATCATC CATATAAATT GAATTTTAAT TCAATTGAGGC GTTAGCCACA GGAGGGATCT																								
ATG	TCC	GGG	TTT	TAT	CAT	AAG	CAT	TTT	CTG	AAA	TTA	CTC	GAT	TTC	ACG	CCA	GCT	GAA	CTC	AAC	AGC	CTG	CTG	CAG	25
met	ser	gly	phe	tyr	his	lys	his	phe	leu	lys	leu	leu	asp	phe	thr	pro	ala	glu	leu	asn	ser	leu	leu	gln	
TTA	GCC	GCG	AAG	CTG	AAA	GCC	GAT	AAG	AAA	GGT	AAA	GAA	GAA	GCC	AAA	CTC	ACT	GGT	AAA	AAC	ATC	GCG	CTC		50
leu	ala	ala	lys	leu	lys	ala	asp	lys	lys	ser	gly	lys	glu	glu	ala	lys	leu	thr	gly	lys	asn	ile	ala	leu	
ATC	TTC	GAA	AAA	GAC	TCG	ACT	CGT	ACC	CGA	TGC	TCT	TTC	GAA	GTT	GCC	GCA	TAT	GAC	CAG	GGT	GCT	GCG	GTT	ACT	75
ile	phe	glu	lys	asp	ser	thr	arg	thr	arg	cys	ser	phe	glu	val	ala	ala	tyr	asp	gln	gly	ala	arg	val	thr	
TAT	CTC	GGC	CCA	AGC	GGC	AGC	CAG	ATT	GGT	CAT	AAA	GAG	TCG	ATT	AAA	GAC	ACT	GCC	GCG	GTG	CTT	GGT	CGC	ATG	100
tyr	leu	gly	pro	ser	gly	ser	gln	ile	gly	his	lys	gly	ser	ile	lys	asp	thr	ala	arg	val	leu	gly	arg	met	
TAT	GAC	GGT	ATT	CAG	TAT	CGC	GGC	TAT	GGT	CAG	GAG	ATT	GTC	GAA	ACA	CTG	GCG	CAA	TAC	CGT	AGC	GTG	CCG	GTA	125
tyr	asp	gly	ile	gln	tyr	arg	gly	tyr	gly	gln	glu	ile	val	glu	thr	leu	ala	gln	tyr	arg	ser	val	pro	val	
TGG	AAT	GGC	CTG	ACC	AAT	GAG	TTC	CAT	CCC	ACG	CAG	CTG	ATC	GAG	TAT	AAA	CTC	ACC	ATG	CAG	GAG	CAT	TTG	CCC	150
trp	asn	gln	thr	asn	glu	phe	his	pro	thr	gln	leu	ile	glu	tyr	lys	leu	thr	met	gln	his	leu	leu	pro		
GGC	AAA	GCG	TTT	CAA	GAA	ATG	ACG	CTG	GTC	TAT	GCA	GGT	GAC	GCG	CGT	AAC	AAC	ATG	GGC	AAT	TCG	ATG	CTC	GAA	175
gly	lys	ala	phe	asn	gln	met	thr	leu	val	tyr	ala	gly	asp	ala	arg	asn	asn	met	gly	asn	ser	met	leu	glu	
GCT	GCG	GCG	CTT	ACC	GGT	CTG	GAT	TTG	CGT	CTG	GTC	GCG	CCA	CAA	GCG	TGC	TGG	CCG	GAA	GCT	GCG	CTG	GTT	ACG	200
ala	ala	ala	leu	thr	gly	leu	asp	leu	arg	leu	val	ala	pro	gln	ala	cys	trp	pro	glu	ala	ala	leu	val	thr	
GAA	TGC	GCG	GCC	CTG	GCA	CAG	CAA	AAT	GGT	GGG	AAT	ATT	ACG	CTG	ACT	GAA	GAT	GTC	GCG	AAG	GGG	GTT	GAA	GGT	225
glu	cys	arg	ala	leu	ala	gln	gln	asn	gly	asn	ile	thr	leu	thr	glu	asp	val	ala	lys	gln	ala	val	gly	gly	
GCT	GAC	TTT	ATC	TAT	ACC	GAT	GTG	TGG	GTG	TCG	ATG	GGG	GAA	GCA	AAA	CAG	AAA	TGG	GCG	GAA	CSG	ATT	GCA	TTG	250
ala	asp	phe	ile	tyr	thr	asp	val	trp	val	ser	met	gly	glu	ala	lys	gln	lys	trp	ala	glu	arg	ile	ala	leu	
CTG	GCT	GAA	TAT	CAG																					

b LOCALLY HOMOLOGOUS REGIONS BETWEEN *argI* & *pyrB*

```

-35          +5
argI      CAAATGAATAATCAT CCATATAAATTGAATTTTAATTCA
          .....
          CAAAT AA AA ATGC ATATACCTTGACTTTTAATTCA
-108       -73

```

Figure 4. a. The DNA sequence of argI and the deduced amino acid sequence of the OTCase polypeptide. The transcriptional control region is presented in italics. The presumed ribosomal binding site (underlined, +25) and promoter region (underlined -10, -35) as reported by Piette et al. (13) are indicated. "+1" indicates the presumed site of transcriptional initiation. Amino acid sequence is numbered on the right side of the figure. b. Comparison of the argI promoter sequence with a nucleotide sequence preceding the pyrBI operon attenuator (3).

Marshall and Cohen have identified a catalytically essential cysteinyl residue which perturbs the binding of ornithine in streptococcal and bovine liver OTCase (22, 23). The cysteines are located at 66 and 51 residues from the C-termini of the respective enzymes. Horwich et al. (24) have also shown that rat liver OTCase

possesses a thiol residue 52 amino acids from its C-terminus. In this study, the deduced polypeptide sequence places a cysteine 61 residues from the C-terminus of OTCase-1 of E. coli. Kuo et al. (25) have demonstrated that ornithine and Zn^{++} compete for the same site on the E. coli enzyme. Thus, it can be proposed that this interaction involves the cysteinyl residue, which might be involved in zinc binding (similar to that of the ATCase regulatory subunit (11) or alcohol dehydrogenase) (26). The thiol groups of the mammalian enzymes are similarly located in the primary sequence (52 and 51 residues) as the bacterial enzymes (66 and 61 residues). These considerations suggest that this particular residue might perform an essential function in the binding of ornithine.

OTCase and ATCase are anabolic enzymes which are often required in relatively high levels in order to meet the cellular needs for arginine and pyrimidines. Enzyme levels in both systems are normally controlled by the rate of transcriptional initiation (3,14,27,28). Preliminary comparisons of the regulatory regions of argI and argF have been reported by Piette et al. (14). The nucleotide sequence of the promoters and presumed operators are highly homologous, as are the regions preceding the control regions. Although these regions show no homology with the attenuator sequence of pyrB (3,27,28), there is a striking homology with a preceding region located 35 to 70 base pairs before the consensus promoter (-73 to -180 relative to the presumed transcriptional start for the pyrBI operon) (3). This sequence is quite similar to the promoter/operator region of argI (see Figure 4b), however, it does not appear to function as a legitimate promoter for pyrBI since in vitro transcriptional studies of pyrBI mRNA (27,28) have identified two different putative promoters. Nonetheless, this homology may represent some heretofore unrecognized, common evolutionary ancestry between the regulatory regions of argI/argF and the pyrI operon. As demonstrated in earlier comparisons (8,14,21) the structural regions of argI, argF and pyrB possess significant homologies in both amino acid sequence and nucleotide sequence. Thus, it appears that the sequence homology of argI and argF with pyrB is broken by the attenuator region of pyrB. Perhaps this indicates the introduction of an attenuator module into some ancestral gene in order to potentiate the expression of pyrB. In order to maximize expression in "highly expressed gene systems," codon usage must be optimized to minimize the number

of interactions of the message with minor tRNA species which could limit the rate of translation (29,30). An analysis of various genes from E. coli has established a strong correlation between codon usage and gene expressivity (30). Consistent with these observations, the codon usage pattern of argI is similar to that of pyrB and other highly expressed genes (D.A. Bencini, 1983, Ph.D. Dissertation, Texas A&M University). The codons AGA/AGG(arg), UCA(ser), CCU(pro), UGU(cys), and AUA(ile) represent minor tRNA species and are not used in the wild-type argI mRNA. Similarly, the pyrB message does not utilize CGA/CGG/AGA/AGG(arg), CCU(pro), and UGC(cys). However, a few codons which correspond to minor tRNA species are occasionally used. The argI message used UGC(ser) and the pyrB message contained AUA(ile) although these codons are infrequently used in messages from highly expressed genes. These results are consistent with the hypothesis that highly expressed genes such as argI contain a relatively low incidence of codons for minor tRNA species which may reduce translational efficiency.

This study reports the complete DNA sequence of the structural gene region of argI from E. coli K-12. The deduced amino acid sequence and nucleotide sequence compare favorably with previously published partial sequences (14,15,21). The availability of this sequence will contribute to broad scale evolutionary comparisons of primary and secondary structure among the carbamoyltransferase enzymes, not only within bacterial species, but also between quite divergent genera.

ACKNOWLEDGEMENT

This investigation was supported by grants from the Robert A. Welch Foundation (Grant No. A-915), The National Institute of General Medical Sciences (Grant No. 1.R01 GM 29152-01), and the National Science Foundation (Grant No. 902 1983). We wish to particularly thank Mary L. Garcia for excellent secretarial assistance and support.

*To whom communications should be addressed

+Present address: Department of Chemistry, Gibbs Chemical Laboratories, Harvard University, Cambridge, MA 02138, USA

REFERENCES

1. Glansdorff, N., Sand, G., and Verhoef, C. (1967) *Mutat. Res.* 4, 743-751.
2. Legrain, C., Halleux, P., Stalon, V., and Glansdorff, N. (1972)

- Eur. J. Biochem. 27, 93-102.
3. Roof, W.D., Foltermann, K.F., and Wild, J.R. (1982) *Mol. Gen. Genet.* 187, 391-400.
4. Sens, D., Natter, W., and James, E. (1977) *Cell* 10, 275-285.
5. Umbarger, H.E. (1978) *Ann. Rev. Biochem.* 47, 533-606.
6. Lee, J.S., An, G., Friesen, J.D., and Fiil, N.P. (1981) *Cell* 25, 251-258.
7. Gerhart, J.C. and Schachman, H.K., (1965) *Biochemistry* 4, 1054-1062.
8. Gigot, D., Glansdorff, N., Legrain, C., Pierard, A., Stalon, D., Konigsberg, W., Caplier, I., Strosberg, A.D., and Herve, G. (1977) *FEBS Letters* 81, 28-32.
9. Honzatko, R.B., Crawford, J.L., Monaco, H.L., Ladner, J.E., Edwards, B.F.P., Evans, D.R., Warren, S.G., Wiley, D.C., Ladner, R.C., and Lipscomb, W.N. (1982) *J. Mol. Biol.* 160, 219-264.
10. Honzatko, R.B., and Lipscomb, W.N. (1982) *J. Mol. Biol.* 160, 265-286.
11. Kantrowitz, E.R., Pastra-Landis, S.C., and Lipscomb, W.N. (1980) *Trends Biochem. Sci.* 5, 124-128.
12. Hoover, T.A., Roof, W.D., Foltermann, W.D., O'Donovan, G.A., Bencini, D.A., and Wild, J.R. (1983) *Proc. Natl. Acad. Sci. USA* 80, 2402-2466.
13. Konigsberg, W.H., and Henderson, L. (1983) *Proc. Natl. Acad. Sci. USA* 80, 2467-2471.
14. Piette, J., Cunin, R., Van Vliet, F., Charlier, D., Crabeel, M., Ota, Y., and Glansdorff, N. (1982) *The EMBO Journal* 1, 853-857.
15. Moore, S.K. (1981) Ph.D. Dissertation, Univ. of Kentucky.
16. Vierira, J., and Messing, J. (1982) *Gene* 19, 259-268.
17. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) In *Molecular Cloning: A Laboratory Manual*, pp. 113-126, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
18. Maxam, A., and Gilbert, W. (1980) *Methods in Enzymology* 65, 499-560.
19. Hansen, J.N. (1981) *Anal. Biochem.* 116, 146-151.
20. Smith, D.R., and Calvo, J.M. (1980) *Nucleic Acid Res.* 8 2255-2274.
21. Piette, J., Gigot, D., Lauwereys, M., Cunin, R., Glansdorff, N., and Pierard, A. (1982) *Archives Internationales de Physiologie et de Biochemie* 90, B211-B212.
22. Marshall, M., and Cohen, P.P. (1972) *J. Biol. Chem.* 247, 1641-1653.
23. Marshall, M., and Cohen, P.P. (1980) *J. Biol. Chem.* 255, 7296-7300.
24. Horwich, A.L., Kraus, J.P., Williams, K., Kalousek, F., Konigsberg, W., and Rosenberg, L.E. (1983) *Proc. Natl. Acad. Sci. (USA)* 80, 4258-4262.
25. Kuo, L.C., Lipscomb, W.N., and Kantrowitz, E.R. (1982) *Proc. Natl. Acad. Sci. USA* 79, 2250-2254.
26. Coleman, J.E., (1979) in *Zinc (a report of the U.S. National Research Council)*, pp. 211-223, University Park Press, Baltimore, Md., USA.
27. Navre, M. and Schachman, H.K. (1983) *Proc. Natl. Acad. Sci. (USA)* 80, 1207-1211.
28. Turnbough, C.L., Hicks, K.L., and Donahue, J.P. (1983) *Proc. Natl. Acad. Sci. (USA)* 80, 368-372.
29. Grosjean, H., and Fiers, W. (1982) *Gene* 18, 199-209.
30. Gouy, M., and Gautier, C. (1982) *Nucleic Acids Res.* 10, 7055-7074.