Volume 215, number 2, 291-294 FEB 04675 May 1987

The nucleotide sequence of the *amiE* gene of *Pseudomonas aeruginosa*

W.J. Brammar, I.G. Charles*, M. Matfield†, Liu Cheng-Pin△, R.E. Drew+ and Patricia H. Clarke+°

Department of Biochemistry, University of Leicester, University Road, Leicester LE1 7RH and +*Department of Biochemistry, University College London, Gower Street, London WC1E 6BT, England*

Received 20 January 1987

The nucleotide sequence of the *amiE* gene, encoding the aliphatic amidase of *Pseudomonas aeruginosa*, has been determined. The sequence of 1038 nucleotides shows a strong bias in favour of codons with G or C in the third position, and only 44 different codons are utilised.

Aliphatic amidase; Coding sequence; Codon usage; *(Pseudomonas aeruginosa)*

1. INTRODUCTION

The utilization of short-chain aliphatic amides by *Pseudomonas aeruginosa* depends on the activity of an inducible aliphatic amidase (EC 3.5.1.4) [1,2]. The structural gene for the enzyme, *amiE*, is capable of high rates of expression under positive control by the product of a linked regulator gene, *amiR* [3–5]. The amidase genes of PAC433, an up-promotor mutant of *Ps. aeruginosa* strain PAC1 [6], have previously been cloned in a bacteriophage vector and expressed in *Escherichia coli* [7].

The molecular mass of the aliphatic amidase determined by sedimentation equilibrium was found to be 200000 Da. The enzyme is a hexamer of identical subunits whose molecular mass, determined by gel electrophoresis and gel filtration, was reported as 33000–35000 Da [8]. The amino acid sequence of the amidase has been substantially determined [9,10]. Mutants with altered substrate-specificities have been isolated [11–13] and two of the amino acid substitutions have been identified [13,14].

We present here the nucleotide sequence of the *amiE* gene of *Ps. aeruginosa*. The accompanying paper [10] documents the amino acid sequence of the amidase protein determined by protein chemistry.

Correspondence address: W.J. Brammar, Dept of Biochemistry, University of Leicester, University Road, Leicester LE1 7RH, England

* Present address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH, England

† Present address: MRC, 20 Park Crescent, London W1N 4AL, England

△ Present address: Department of Biochemistry, Louisiana State University, Baton Rouge, LA 70803, USA

° Present address: Department of Chemical and Biochemical Engineering, University College, London, WC1E 7JE, England

2. MATERIALS AND METHODS

2.1. *Materials*

Deoxy- and dideoxy-nucleoside triphosphates, ampicillin, dextran sulphate and dithiothreitol were from Sigma; restriction endonucleases were from BRL (Gibco) or Pharmacia, T_4 DNA ligase was from Boehringer and large fragment (Klenow) *E. coli* DNA polymerase I was from Pharmacia. All other reagents were of Analar or greater purity and obtained from local suppliers.

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

2.2. *DNA preparations*

Plasmid DNAs were prepared from *E. coli* by detergent lysis and CsCl-ethidium bromide equilibrium centrifugation [15]. Bacteriophage M13 RF DNA was isolated from infected cultures of *E. coli* strain JM101 [16], using the method of Birnboim and Doly [17].

2.3. *Preparation of sequencing templates*

Required fragments of DNA for sequencing were separated by electrophoresis on low melting-point agarose gels, excised and purified by extraction with hot phenol, recovered by precipitation with ethanol and ligated to M13mp8 or M13mp9 vector DNA [18] by conventional methods [19]. Recombinant DNAs were recovered as phage in colourless M13 plaques following transfection of *E. coli* strain JM101 and single-stranded DNAs for use as templates were purified by standard methods [20].

2.4. *DNA sequencing*

The DNA sequencing methods using ^{35}S-dATP, dideoxynucleoside triphosphates and buffer-gradient gels were used as previously described [20]. Universal primer and oligodeoxynucleotides as internal sequencing primers were prepared by John Keyte by the method of Matthes et al. [21] with modifications to the wash-cycle as described by Sproat and Gait [22]. The strategy for sequencing the *amiE* gene is shown in fig.1. Difficulties were sometimes encountered in obtaining readable sequence data in specific regions. These were usually overcome by carrying out the polymerase reaction at 37°C or by the use of sequencing gels containing 20% formamide.

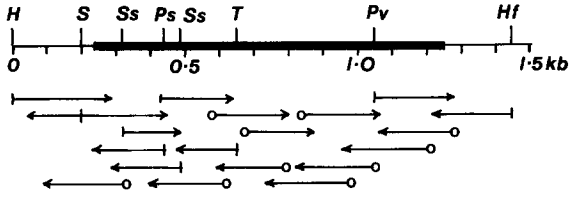

Fig.1. Strategy for sequencing the *amiE* gene. The map shows the positions of some restriction enzyme sites flanking and within the gene. (The *amiE* coding sequence is shown as the thickened line.) The horizontal arrows beneath the map show the directions and extents of sequencing reactions using universal primer to sequence from known restriction sites (↦) or specific oligonucleotides as internal primers (○→). The restriction enzyme sites are: *Hin*dIII (H); *Sma*I (S); *Sst*II (Ss); *Pst*I (Ps); *Taq*I (T); *Pvu*II (Pv) and *Hin*fI (Hf).

3. RESULTS AND DISCUSSION

The nucleotide sequence encoding the aliphatic amidase of *Ps. aeruginosa* strain PAC433 is shown

```
      10                  30                  50
AAATAACAACAGGAGGTGATACCCATGCGTCACGGCGATATTTCCAGCAGCAACGACACC
                    MetArgHisGlyAspIleSerSer ArgAsnAspThr
      70                  90                 110
GTCGGAGTGGCCGTGGTCAACTACAAGATGCCGCGCCTGCACACCGCGA GGAGGTCCTG
ValGlyValAlaValValAsnTyrLysMetProArgLeuHisThrAlaArgGluValLeu
     130                 150                 170
GACAACGCCCGGAAGATCGCCGACATGATCGTCGGCATGAAGCAGGGC TGCCGGCATG
AspAsnAlaArgLysIleAlaAspMetIleValGlyMetLysGlnGlyLeuProGlyMet
     190                 210                 230
GACCTGGTGGTGTTCCCGGAGTACAGCCTGCAGGGCATCATGTACGA GCGGCGGAGATG
AspLeuValValPheProGluTyrSerLeuGlnGlyIleMetTyrAspProAlaGluMet
     250                 270                 290
ATGGAAACCGCGGTGGCGATCCCCGGCGAGGAAACCGAGATATTCTCCCGCGCCTGCCGC
MetGluThrAlaValAlaIleProGlyGluGluThrGluIlePheSerArgAlaCysArg
     310                 330                 350
AAGGCCAACGTCTGGGGCGTATTCTCCCTCACCGGCGAACGGCACGAGGAGCATCCGCGC
LysAlaAsnValTrpGlyValPheSerLeuThrGlyGluArgHisGluGluHisProArg
     370                 390                 410
AAGGCGCCGTACAACACCCTGGTGCTGATCGACAACAACGGCGAGATC GTCCAGAAGTAC
LysAlaProTyrAsnThrLeuValLeuIleAspAsnAsnGlyGluIleValGlnLysTyr
     430                 450                 470
CGCAAGATCATTCCCTGGTGCCCCATCGAGGGCTGGTATCCCGGTGGCCAGACCTACGTC
ArgLysIleIleProTrpCysProIleGluGlyTrpTyrProGlyGlyGlnThrTyrVal
     490                 510                 530
AGCGAAGGGCCGAAGGGCATGAAGATCAGCCTGATCATCTGCGACGACGG AATTACCCG
SerGluGlyProLysGlyMetLysIleSerLeuIleIleCysAspAspGlyAsnTyrPro
     550                 570                 590
GAGATCTGGCGCGACTGCGCGATGAAGGGCGCCGAGCTGATCGTGCGCTGCCAGGGCTAC
GluIleTrpArgAspCysAlaMetLysGlyAlaGluLeuIleValArgCysGlnGlyTyr
     610                 630                 650
ATGTACCCGGCCAAGGACCAGCAGGTGATGATGGCCAAGGCCATGGCCTGGGCCAACAAC
MetTyrProAlaLysAspGlnGlnValMetMetAlaLysAlaMetAlaTrpAlaAsnAsn
     670                 690                 710
TGCTATGTGGCGGTGGCCAACGCGGCCGGCTTCGACGGTGTCTATTCCTACTTCGGCCAC
CysTyrValAlaValAlaAsnAlaAlaGlyPheAspGlyValTyrSerTyrPheGlyHis
     730                 750                 770
TCGGCGATCATCGGCTTCGACGGCCGTACCCTCGGTGAGTGCGGCGAGGAGGAAATGGGT
SerAlaIleIleGlyPheAspGlyArgThrLeuGlyGluCysGlyGluGluGluMetGly
     790                 810                 830
ATCCAGTACGCCCAGCTGTCCCTTTCGCAGATCCGCGATGCGCGCGCCAACGATCAGTCG
IleGlnTyrAlaGlnLeuSerLeuSerGlnIleArgAspAlaArgAlaAsnAspGlnSer
     850                 870                 890
CAGAACCACCTGTTCAAGATCCTCCACCGCGGCTACAGCGGCTTGCAGGCGTCCGGCGAC
GlnAsnHisLeuPheLysIleLeuHisArgGlyTyrSerGlyLeuGlnAlaSerGlyAsp
     910                 930                 950
GGCGACCGGGGCCTGGCGGAGTGTCCGTTCGAGTTCTACCGCACCTGGGTCACCGACGCC
GlyAspArgGlyLeuAlaGluCysProPheGluPheTyrArgThrTrpValThrAspAla
     970                 990                1010
GAGAAGGCGCGCGACAATGTCGAGCGACTGACCCGCTCGACCACCGGCGTGGCGCAATGC
GluLysAlaArgAspAsnValGluArgLeuThrArgSerThrThrGlyValAlaGlnCys
    1030                1050
CCGGTCGGCCGGCTGCCCTATGAGGGACTGGAGAAGGAAGCCTGA
ProValGlyArgLeuProTyrGluGlyLeuGluLysGluAlaEnd
```

Fig.2. The nucleotide sequence of the coding region of the *amiE* gene. The DNA strand shown has the same orientation as the messenger RNA. The nucleotides are numbered from an arbitrary start-point that precedes the putative ribosome-binding site.

in fig.2. An open reading-frame of 1038 nucleotides is initiated by an ATG codon and completed by a TGA translation-termination codon. The protein encoded by this nucleic acid sequence contains 346 amino acids and has a molecular mass of 38378 Da. Starting 14 nucleotides upstream of the ATG initiation codon is a pentanucleotide sequence (-AGGAG-) that is complementary to five nucleotides near the 3′-end of *Ps. aeruginosa* 16 S ribosomal RNA [23], and which presumably constitutes the ribosome-binding site. This same sequence is part of a decanucleotide (-AGGAGGT GAT-) that has perfect complementarity with a sequence near the 3′-end of *E. coli* 16 S rRNA [24]. As noted by Atkins [25] for *E. coli* sequences, UAA and UGA termination codons are found closely preceding the translation-initiation codon.

The codon-usage within the *amiE* gene is shown in table 1. The bias in favour of G or C in the third position of the codons is striking: of 346 codons, only 33 (9.6%) have A or T in the third position. Seventeen of the possible sense codons are not used in the gene, with bias against codons rich in A or T. Thus all 20 arginine codons are the CGX type, AG not being used; 19 leucine codons are of the form CTX, the only other one being TTG. The only codon used to encode the 13 threonine residues is ACC, while AAG is used exclusively to encode the 15 lysines. It is not the case that *Pseudomonas* species are incapable of utilising certain codons because of the absence of the cognate transfer RNAs, since the entire codon repertoire is used in genes such as *xylS* of *Pseudomonas putida* mt-2 [26]. The highly biased codon usage appears to be characteristic of genes that are expressible at high rates, including the mercuric reductase gene of Tn501 [27], the *xylE* gene on the Tol plasmid [28] and the carboxypeptidase G2 gene [29].

The complete sequence of the aliphatic amidase now allows the amino acid replacements found in substrate-specificity variants to be placed within the primary structure. The B6 mutation, a substitution of phenylalanine for serine that potentiates hydrolysis of butyramide, involves a C → transition within codon 7 [14]. The A13 mutation, a threonine to isoleucine change that allows utilisation of acetanilide [13], involves an ACC → AUC change in the codon for *thr*103. In the absence of information about the tertiary structure of the protein, we cannot tell whether the amino acids at positions 7 and 103 are involved in the substrate-binding site, or whether they simply influence substrate-binding at a distance by more general effects on enzyme-conformation. Knowledge of the coding sequence of the *amiE* gene should facilitate the determination of the nature of the mutational events giving rise to other variants, via cloning and sequencing using a set of internal primers.

Table 1

Codon-usage within the *amiE* gene

Phe	UUU	0	Ser	UCU	0	Tyr	UAU	4	Cys	UGU	1
	UUC	9		UCC	6		UAC	13		UGC	8
Leu	UUA	0		UCA	0	Ter	UAA	0	Ter	UGA	1
	UUG	1		UCG	4	Ter	UAG	0	Trp	UGG	6
Leu	CUU	1	Pro	CCU	0	His	CAU	1	Arg	CGU	2
	CUC	3		CCC	6		CAC	6		CGC	13
	CUA	0		CCA	0	Gln	CAA	1		CGA	1
	CUG	15		CCG	10		CAG	13		CGG	4
Ile	AUU	2	Thr	ACU	0	Asn	AAU	2	Ser	AGU	0
	AUC	18		ACC	13		AAC	12		AGC	6
	AUA	1		ACA	0	Lys	AAA	0	Arg	AGA	0
Met	AUG	15		ACG	0		AAG	15		AGG	0
Val	GUU	0	Ala	GCU	0	Asp	GAU	4	Gly	GGU	4
	GUC	11		GCC	17		GAC	15		GGC	27
	GUA	1		GCA	0	Glu	GGA	6		GGA	2
	GUG	11		GCG	15		GAG	20		GGG	1

ACKNOWLEDGEMENTS

This paper is dedicated to the memory of Judy Bundick, who gave valuable technical assistance in the early stages of the project. We are grateful to John Keyte for oligodeoxyribonucleotides, to Dr Richard Ambler for the exchange of unpublished information and to Anne Bates for typing the manuscript. The work was supported by a project grant from the Medical Research Council.

REFERENCES

[1] Kelly, M. and Clarke, P.H. (1962) J. Gen. Microbiol. 27, 305–316.
[2] Brammar, W.J. and Clarke, P.H. (1964) J. Gen. Microbiol. 37, 307–319.
[3] Brammar, W.J., Clarke, P.H. and Skinner, A.J. (1967) J. Gen. Microbiol. 47, 87–102.
[4] Farin, F. and Clarke, P.H. (1978) J. Bacteriol. 135, 379–392.
[5] Drew, R.E. (1984) J. Gen. Microbiol. 130, 3101–3111.
[6] Smyth, P.F. and Clarke, P.H. (1975) J. Gen. Microbiol. 90, 91–99.
[7] Drew, R.E., Clarke, P.H. and Brammar, W.J. (1980) Mol. Gen. Genet. 177, 311–320.
[8] Brown, P.R., Smyth, M.J., Clarke, P.H. and Rosemeyer, M.A. (1973) Eur. J. Biochem. 34, 177–187.
[9] Auffret, A.D. (1977) PhD Thesis, University of Edinburgh.
[10] Auffret, A.D., Ambler, R.P. and Clarke, P.H. (1987) FEBS Lett. 215, 285–290.
[11] Brown, J.E., Brown, P.R. and Clarke, P.H. (1969) J. Gen. Microbiol. 57, 273–295.
[12] Betz, J.L. and Clarke, P.H. (1972) J. Gen. Microbiol. 73, 161–174.
[13] Brown, P.R. and Clarke, P.H. (1972) J. Gen. Microbiol. 70, 287–298.
[14] Paterson, A. and Clarke, P.H. (1979) J. Gen. Microbiol. 114, 75–85.
[15] Clewell, D. and Helinski, D. (1969) Proc. Natl. Acad. Sci. USA 62, 1159–1166.
[16] Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene 33, 103–119.
[17] Birnboim, H.L. and Doly, J. (1979) Nucleic Acids Res. 7, 1513–1523.
[18] Norrander, J., Kempe, T. and Messing, J. (1983) Gene 26, 101–106.
[19] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in: Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory, New York.
[20] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[21] Matthes, M.W.D., Zenke, W.M., Grundstrom, T., Staub, A., Wintzerith, M. and Chambon, P. (1984) EMBO J. 3, 801–805.
[22] Sproat, B.S. and Gait, M.J. (1985) Nucleic Acids Res. 13, 2959–2978.
[23] Shine, J. and Dalgarno, L. (1975) Nature 254, 34–38.
[24] Shine, J. and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1342–1346.
[25] Atkins, J.F. (1979) Nucleic Acids Res. 7, 1035–1041.
[26] Spooner, R.A., Lindsay, K. and Franklin, F.C.H. (1986) J. Gen. Microbiol. 132, 1347–1358.
[27] Brown, N.L., Ford, S.J., Pridmore, R.D. and Fritzinger, D. (1983) Biochemistry 22, 4089–4095.
[28] Nakai, C., Kagamiyama, H., Nozaki, M., Nazawa, T., Inouye, S., Ebina, Y. and Nakazawa, A. (1983) J. Biol. Chem. 258, 2923–2928.
[29] Minton, N.P., Atkinson, T., Bruton, C.J. and Sherwood, R.F. (1984) Gene 31, 31–38.