Nucleotide sequence of a variant antifreeze protein gene

Sherry Gauthier, Yaling Wu and Peter L.Davies*

Department of Biochemistry, Queen's University, Kingston, Ontario K7L 3N6, Canada

Submitted July 18, 1990

EMBL accession no. X53718

Antifreeze protein (AFP) genes from the winter flounder have been classified by their organization as either tandemly repeated, or linked but irregularly spaced (1). We report here the nucleotide sequence of an AFP gene belonging to the latter category. The gene was present in clone $\lambda 2-1$ isolated from a λ Charon 4 genomic library and was subcloned into pBR322 to give plasmid 21a as previously described (2). The gene sequence shown in Fig. 1 corresponds to one of two regions in 21a that hybridize to AFP cDNA, the one closest to the BamHI site. Based on its nucleotide sequence the gene is similar in organization to two other AFP genes previously reported from the winter flounder (2, 3), but differs by encoding a 91 amino acid preproprotein that contains four 11-amino acid repeats in the mature protein rather than the usual three. A protein matching this composition and sequence does not appear to be a major component of the circulating AFP in winter flounder (4). However, a very similar cDNA sequence (clone 11A7) has been reported which differs in only three bases, one at 1022 in the 3' untranslated region, the others at 920 and 921 in the 3' exon (5). The latter changes would cause a Ala \rightarrow Val substitution.

ACKNOWLEDGEMENT

This work was supported by a grant from the Medical Research Council of Canada.

REFERENCES

- Scott,G.K., Hew,C.L. and Davies,P.L. (1985) Proc. Natl. Acad. Sci. USA 82, 2613-2617.
- Davies, P.L., Hough, C., Scott, G., Ng, N., White, B.N. and Hew, C.L. (1984) J. Biol. Chem. 259, 9241-9247.
- Scott,G.K., Davies,P.L., Kao,M.H. and Fletcher,G.L. (1988) J. Mol. Evol. 27, 29–35.
- Fourney, R.M., Joshi, S.B., Kao, M.H. and Hew, C.L. (1984) Can. J. Zool. 62, 28-33.
- Gourlie, B., Lin, Y., Price, J., DeVries, A.L., Powers, D. and Huang, R.C.C. (1984) J. Biol. Chem. 259, 14960-14965.

GCACAACACTGGGGGGAGTGTTGTACCAATCTGCTCAGATTGGTCGACAGTCAAGCGATGA CCCAGGCTCCAGTTACTATAAAACAGATTCACATTGACCTGGATATTCACCACATGTCA	60 120
MetAlaLeuSerLeuPheThrVa	
TTTTGTAGTAAACCAGTGCTCCCTACAAGTTCTCAAAATGGCTCTCTCACTTTTCACTGT	180
lGlyGlnLeuIlePheLeuPheTrpThrMetArg	
CGGACAATTGATTTTCTTATTTTGGACAATGAGgtacgtgaacactcactttgtttcttc	240
concentration in the second	240
tatgaatctggttttactgtaaatatcttgggaaggaagg	300
cgaggggccatttgctttacagccagcggtgaaagatgaagatcttcatctgtgttcatc	360
tgatggaaagttagttctgaaaccttcattggaagaatcagattcatgtgttcaggctta	420
aaggacgtttttagactcgtctgttaaatcatgggaaacaactttttaattcagtcag	480
ctggaaaactcttttatatgcacagaagaagaagaagtgatctttagttcatcaccatgg	540
aaacatcatcagcagttaaagtttgtctgcttcagtatcaccggccagttccagtgctca	600
${\tt atgattttagtttgaatgaagaagctgtcatttggttttatgttgggggggg$	660
ggggggggtcatcacacacagatattgaaaactgtcatcactgggtctggtgaaagtgacg	720
gaccagtaaatgttgtgataataatattatcacaataattataataataccattaatctc	780
IleThrGluAlaAsnProAspProAlaAlaLysAlaValProAlaAlaAlaAla	
tgcagAATCACTGAAGCCAACCCCGACCCCGCAGCCAÄAGCCGTCCCTGCCGCAGCCGCC	840
↓ MATURE	
ProAspThrAlaSerAspAlaAlaAlaAlaAlaAlaAlaAlaThrAlaAlaThrAlaAlaAla	
CCAGACACCGCCTCTGATGCCGCCGCAGCAGCCGCCGCCACCGCAGCCACCGCCGCCG	900
¢	200
AlaAlaAlaAlaThrAlaValThrAlaAlaLysAlaAlaAlaLeuThrAlaAlaAsnAla	
GCAGCCGCCGCCACCGCAGTCACCGCTGCCAAAGCCGCAGCCCTCACCGCCGCCAACGCC	960
¥¥	
3]=3]=3]=3]=3]=3]=mbus]=3]=3]=3]=3]=3]=3]=3]=3]=3]	
AlaAlaAlaAlaAlaAlaThrAlaAlaAlaAlaAlaArgGlyEND GCCGCCGCCGCAGCAGCCACCGCCGCCGCAGCCGCCAGAGGTTAAGGATCGTCGTCGTCT	1020
	1020
•	
TGATGTGGGATCATGTGAACATCTGAGCAGCGAGATGTTACCAATCTGTTG AATAAA GCT	1080
	1140
GTTTCTTTAGATCATGTTTTTCTTAACACATAAACCTCCAGAAATCATGATGCGTCACGT	1200
	1260
CTTGTTTCCATTATGCTAGGCTAACAGACGACGGACTGCAG	1301

Figure 1. AFP gene sequence. Intron sequence is in lower case. Putative CAAT and TATA boxes, the initiating methionine codon and the polyadenylation signal are highlighted. The predicted start of the mature protein sequence is shown by the vertical arrow, and the four 11-amino acid repeats are underlined.