Volume 251, number 1,2, 197–200 FEB 07386 July 1989

Isolation of a cDNA encoding the human G_{M2} activator protein

M. Schröder, H. Klima, T. Nakano*, H. Kwon, L.E. Quintern, S. Gärtner, K. Suzuki* and K. Sandhoff

Institut für Organische Chemie und Biochemie, Gerhard-Domagk Str. 1, 53 Bonn 1, FRG and **Biological Sciences Research Center, Departments of Neurology and Psychiatry, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*

Received 24 May 1989

The G_{M2} activator protein is a glycolipid-binding protein required for the lysosomal degradation of ganglioside G_{M2}. A human fibroblast cDNA library was screened with mixtures of oligonucleotide probes corresponding to four different areas of the amino acid sequence. A putative clone (821 bp) which gave positive signals to all four probe mixtures was purified and sequenced. The sequence was colinear with the sequence of 160 amino acids of the mature G_{M2} activator protein. Availability of the cDNA clone should facilitate investigation into function of the G_{M2} activator protein and also into genetic abnormalities underlying G_{M2} gangliosidosis AB variant.

G_{M2} activator protein; G_{M2} gangliosidosis; cDNA sequence

1. INTRODUCTION

Hydrolysis of ganglioside G_{M2} requires G_{M2} activator protein in addition to the lysosomal β-hexosaminidase A, which consists of two subunits, α and β. Inherited defects in any of the three protein components in the system, β-hexosaminidase α- and β-subunits and the G_{M2} activator protein thus cause a G_{M2} gangliosidosis [1]. G_{M2} gangliosidosis due to defective G_{M2} activator protein but with normal β-hexosaminidase subunits is known as AB variant [1]. cDNA and genomic clones of both of the β-hexosaminidase have been isolated and characterized, and a dozen genetic abnormalities causing the disease have already been identified. However, G_{M2} activator protein has not been cloned. In this report we describe isolation of a cDNA clone that encodes the mature normal human G_{M2} activator protein.

Correspondence address: K. Sandhoff, Institut für Organische Chemie und Biochemie, Gerhard-Domagk Str. 1, 53 Bonn 1, FRG

2. MATERIALS AND METHODS

An SV40-transformed human fibroblast cDNA library in the Okayama-Berg vector [2] was kindly supplied by Dr H. Okayama. Restriction enzymes and M13 vector were purchased from Pharmacia and Boehringer Mannheim. The sequencing kit and T4 polynucleotide kinase were from Pharmacia. [γ-^{32}P]ATP (5000 Ci/nmol) and [α-^{35}S]dATP (1000 Ci/nmol) were from Amersham Buchler. Nitrocellulose membranes (type HATF) were from Millipore; zeta bind nylon membranes were from BioRad Labs.

2.1. *Oligonucleotide mixtures*

Mixtures of oligonucleotide probes corresponding to amino acid sequences of four different regions of the mature human G_{M2} activator protein [3,4] were synthesized (table 1). They were analyzed by electrophoresis in 20% polyacrylamide/7 M urea gel. The area of the gel containing the oligonucleotide band was cut and DNA eluted overnight. The probes were labelled at the 5'-end with [γ-^{32}P]ATP (5000 Ci/nmol) and T4 polynucleotide kinase [5].

2.2. *cDNA library screening*

The pcD human fibroblast cDNA library [2] was plated at a density of approximately 4000 colonies/134 mm Petri dish. A total of 2×10^5 clones were screened with the oligonucleotide mixture no.3 as the probe (table 1). Positive clones were identified and purified by repeated screening [5]. Prehybridization (6 × SSPE, 10 × Denhardt's solution, 0.5% SDS) and

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

hybridization (6 × SSPE, 5 × Denhardt's solution, 0.5% SDS, 50 pmol of ^{32}P-labelled oligonucleotides) were performed in plastic pouches with 10 filters/bag at 8°C below the lowest T_m for the probe mixture. Filters were washed at the lowest T_m for the oligonucleotide probe mixture in 6 × SSC, 0.1% SDS for 3–4 h.

2.3. *DNA sequencing*

The cDNA insert of the purified clone was subcloned into the M13 vector and sequenced by the dideoxy chain termination method [6] with [α-^{35}S]dATP (1000 Ci/nmol) [7]. The sequence analysis was done independently in the two laboratories and both strands were used as templates throughout.

3. RESULTS AND DISCUSSION

The screening of 2 × 10^5 independent clones with the oligonucleotide mixture 3 (table 1) initially yielded 75 positive clones. Sixty clones were left after the second round of screening. They were characterized by restriction analysis, agarose gel electrophoresis and Southern blotting analysis. One clone (pGAP1) was obtained which was 821 bp long and hybridized to all of the four oligonucleotide probe mixtures.

The sequence of pGAP1 (fig.1) included a segment that was colinear with the amino acid sequence of the mature G_{M2} activator protein as determined by Edman analysis. Ser^{18} is the N-terminal amino acid of the mature G_{M2} activator protein. The cDNA sequence included 17 additional amino acids on the N-terminal side without reaching the initiation codon. Thus, we are unable to determine how large the precursor polypeptide is. On the other hand, the stop codon is present exactly at the location predicted from the amino acid sequence analysis. Three discrepancies were found between the sequences deduced from the nucleotide sequences and that determined by Edman analysis; Val^{55} was where Edman analysis expected a methionine, and the last two amino acids, Gly^{178} and Ile^{179}, were found instead of methionine and serine as predicted by Edman analysis. One putative *N*-glycosylation site is present, Asn-Val-Thr (49–51), in agreement with the Edman analysis. Neither a poly(A)$^+$ tail nor a polyadenylation signal was present in the 3′-untranslated region, indicating that the mRNA is also longer on the 3′-side than pGAP1.

pGAP1 does not represent the full-length mRNA. The mRNA coding for the G_{M2} activator protein must be longer in both 5′- and 3′-directions. Attempts are under way to obtain full-length cDNA clones for definitive characterization. Nevertheless, pGAP1 includes the entire coding sequence of the mature normal human G_{M2} activator protein and thus is useful for analysis of the genomic structure and analysis of specimens from patients with G_{M2} gangliosidosis AB variant [8].

Table 1

Peptide sequences of G_{M2} activator protein and corresponding oligonucleotide probe mixtures

Oligonucleotide	Sequence							
		W	D	N	C	D	E	G
Oligo 1		TGG	GAT	AAT	TGT	GAT	GAA	GG
(20-mer; 32 mix)			C	C	C	C	G	
		F	E	H	F	C	D	V
Oligo 2		TTT	GAA	CAT	TTT	TGT	GAT	GT
(20-mer; 64 mix)		C	G	C	C	C	C	
	C	H	C	P	F	K	E	G
Oligo 3	TGT	CAT	TGT	CCN	TTT	AAA	GAA	GG
(23-mer; 256 mix)	C	C	C		C	G	G	
			C	I	K	I	A	A
Oligo 4			TGT	ATT	AAA	ATT	GCN	GC
(17-mer; 144 mix)			C	C	G	C		
				A		A		

```
                                                  ..... TTG
                                                        L

CTTCTCGCGACCCCTGCGCAAGCCCACCTGAAAAAGCCATCCCAGCTCAGTAGCTTTTCC  63
 L  L  A  T  P  A  Q  A  H  L  K  K  P  S  Q  L  S  S  F  S   21

TGGGATAACTGTGATGAAGGGAAGGACCCTGCGGTGATCAGAAGCCTGACTCTGGAGCCT 123
 W  D  N  C  D  E  G  K  D  P  A  V  I  R  S  L  T  L  E  P   41

GACCCCATCGTCGTTCCTGGAAATGTGACCCTCAGTGTCGTGGGCAGCACCAGTGTCCCC 183
 D  P  I  V  V  P  G  N  V  T  L  S  V  V  G  S  T  S  V  P   61

CTGAGTTCTCCTCTGAAGGTGGATTTAGTTTTGGAGAAGGAGGTGGCTGGCCTCTGGATC 243
 L  S  S  P  L  K  V  D  L  V  L  E  K  E  V  A  G  L  W  I   81

AAGATCCCATGCACAGACTACATTGGCAGCTGTACCTTTGAACACTTCTGTGATGTGCTT 303
 K  I  P  C  T  D  Y  I  G  S  C  T  F  E  H  F  C  D  V  L  101

GACATGTTAATTCCTACTGGGGAGCCCTGCCCAGAGCCCCTGCGTACCTATGGGCTTCCT 363
 D  M  L  I  P  T  G  E  P  C  P  E  P  L  R  T  Y  G  L  P  121

TGCCACTGTCCCTTCAAAGAAGGAACCTACTCACTGCCCAAGAGCGAATTCGTTGTGCCT 423
 C  H  C  P  F  K  E  G  T  Y  S  L  P  K  S  E  F  V  V  P  141

GACCTGGAGCTGCCCAGTTGGCTCACCACCGGGAACTACCGCATAGAGAGCGTCCTGAGC 483
 D  L  E  L  P  S  W  L  T  T  G  N  Y  R  I  E  S  V  L  S  161

AGCAGTGGGAAGCGTCTGGGCTGCATCAAGATCGCTGCCTCTCTAAAGGGCATATAGCAT 543
 S  S  G  K  R  L  G  C  I  K  I  A  A  S  L  K  G  I  *      179

GGCATCTGCCACAGCAGAATGGAGCGGTGTGAGGAAGGTCCCTTTTCCTCTGTTTTGTGT 603

TTGCCAAGGCCAAACTCCCACTCTCTGCCCCCCTTTAATCCCCTTTCTACAGTGAGTCCA 663

CTACCCTCACTGAAAATCATTTTGTACCACTTACATTTTAGGCTGGGGCAAGCAGCCCTG 723

ACCTAAGGGAGAATGAGTTGGACAGTTCTTGATAGCCCAGGGCATCTGCTGGGCTGACCA 783

CGTTACTCATCCCCGTTAACATTCTCTCTAAAGAGCCT                        821
```

Fig.1. Nucleotide and amino acid sequence of pGAP1. The insert of the clone pGAP1 was 821 nucleotides long. The sequences which are colinear with the amino acid sequence used for oligonucleotide preparation are underlined. Amino acid residues which differ from the sequence obtained by protein analysis of the G_{M2} activator protein are double underlined.

Acknowledgements: The authors wish to thank Dr Okayama, the National Institutes of Health, for the human fibroblast pcD library. This work was supported by grants from the Deutsche Forschungsgemeinschaft (Sa 257/13-1), the Fonds der Chemischen Industrie, and NS-24289 and HD-03110 from the Public Health Service, USA.

REFERENCES

[1] Sandhoff, K., Conzelmann, E., Neufeld, E.F., Kaback, M.M. and Suzuki, M. (1989) in: The Metabolic Basis of Inherited Disease, 6th edn (Scriver, C.R. et al. eds) McGraw-Hill, New York, in press.
[2] Okayama, H. and Berg, P. (1983) Mol. Cell. Biol. 3, 280–289.
[3] Conzelmann, E. and Sandhoff, K. (1987) Methods Enzymol. 138, 792–815.
[4] Fürst, W. (1989) PhD Thesis, Rhein. Friedr. Wilh. Universität, Bonn, FRG.
[5] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[6] Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[7] Biggin, M.D., Gibson, J.J. and Hong, G.F. (1983) Proc. Natl. Acad. Sci. USA 80, 3963–3965.
[8] Burg, J., Banerjee, A. and Sandhoff, K. (1985) Biol. Chem. Hoppe-Seyler 366, 887–891.