The EMBO Journal vol.8 no.9 pp.2469-2473, 1989

# Isolation of cDNA clones encoding human acid sphingomyelinase: occurrence of alternatively processed transcripts

Lothar E.Quintern, Edward H.Schuchman[1], Orna Levran[1], Mariko Suchi[1], Klaus Ferlinz, Heinz Reinke[2], Konrad Sandhoff and Robert J.Desnick[1]

Institut für Organische Chemie und Biochemie der Universitat Bonn, Gerhard-Domagk Strasse 1, D-5300, Bonn 1, FRG, [1]Division of Medical and Molecular Genetics, Mount Sinai School of Medicine, New York, NY 10029, USA and [2]Institut für Biochemie der Universität Bielefeld, Universitätsstrasse 48, Bielefeld, FRG

Communicated by W.Neupert

**Acid sphingomyelinase (sphingomyelin phosphodiesterase, EC 3.1.4.12) was purified from human urine and 12 tryptic peptides were microsequenced (128 residues). Based on regions of minimal codon redundancy, four oligonucleotide mixtures were synthesized and oligonucleotide mixture 1 (20mer; 256 mix) was used to screen $3 \times 10^6$ independent recombinants from a human fibroblast cDNA library. Putative positive clones (92) were purified and analyzed by Southern hybridization with oligonucleotide mixtures 2-4. These studies revealed two groups of clones; group I (80 clones; inserts ranging from ~1.2 to 1.6 kb) hybridized with oligonucleotide mixtures 1-4, while group II (12 clones; inserts ranging from ~1.2 to 1.4 kb) hybridized with oligonucleotide mixtures 1-3. Several group II clones had larger inserts than those in group I, but did not hybridize with oligonucleotide mixture 4. Screening of a human placental cDNA library with a 450 bp group I fragment, also resulted in the isolation of group I and II clones. Representative clones from group I (pASM-1) and group II (pASM-2) were sequenced. pASM-1 contained a 1879 bp insert which was colinear with 96 microsequenced amino acids, while the pASM-2 1382 bp insert was colinear with 78 microsequenced residues. Notably, pASM-2 did not have an internal 172 bp sequence encoding 57 amino acid residues, but had instead an in-frame 40 bp sequence encoding 13 amino acids which was not present in pASM-1. These findings demonstrate the presence of two distinct acid sphingomyelinase transcripts in human fibroblasts and placenta and suggest the occurrence of alternative processing of the mRNA encoding this lysosomal hydrolase.**

*Key words:* alternative splicing/lysosomal hydrolase/lysosomal storage disease/Niemann-Pick disease

## Introduction

In most mammalian tissues there are two phosphodiesterases that convert sphingomyelin to ceramide by removing the phosphocholine moiety (Levade *et al.*, 1986). Acid sphingomyelinase (sphingomyelin phosphodiesterase, EC 3.1.4.12; ASM) is a lysosomal enzyme with a pH optimum of ~4.5, whereas neutral sphingomyelinase (NSM) is a plasma membrane enzyme with a neutral pH optimum (~7.0) and a requirement for magnesium (Schneider and Kennedy, 1967). ASM also has phospholipase C activities towards 1,2-diacylglycerophosphocholine and 1,2-diacylglycerophosphoglycerol. Since these enzymes have been difficult to purify, limited information is available for comparison of their biosynthesis, post-translational processing and physical and kinetic properties. Recently, human urinary ASM was purified ~30 000-fold and shown to be a monomeric protein with a mol. wt of ~72 kd (Quintern *et al.*, 1987). The purified enzyme had both sphingomyelinase and phospholipase C activity towards the respective substrates. A recent report has tentatively assigned the gene for ASM to human chromosome 17 based on studies in somatic cell hybrids. (Konrad and Wilson, 1987), indicating that a single polypeptide is responsible for ASM activity.

The inherited deficiency of ASM activity and the resultant lysosomal accumulation of sphingomyelin has been described in several species (Brady, 1983; Kitagawa, 1987). In man, this lysosomal storage disorder, known as Niemann-Pick disease (NP), occurs in two clinically distinct forms, both inherited as autosomal recessive traits. Niemann-Pick A disease (NP-A) is a severe neurologic disorder of infancy characterized by a rapid neurodegenerative course, hepatosplenomegaly, and death by 2-3 years of life. Affected NP-A homozygotes present with failure to thrive and are usually diagnosed during the first 3-6 months of life. In contrast, Niemann-Pick B disease (NP-B) does not have neurologic involvement, but is a reticuloendotheliosis resulting from the progressive lysosomal deposition of sphingomyelin, primarily in bone marrow, liver, spleen and lungs. NP-B is typically diagnosed in childhood and affected homozygotes may expire from respiratory disease complications in the second to fourth decade of life. Complementation studies performed by fusing NP-A and NP-B fibroblasts have shown that the mutations causing ASM deficiency in each subtype were allelic, since the enzymatic defect was not corrected in the somatic cell hybrids (Besley *et al.*, 1980). However, the biochemical and molecular nature of the clinical heterogeneity resulting from deficient ASM activity remains an enigma. In addition, the fact that homozygotes with either NP-A or NP-B subtypes have normal NSM activities indicated that these enzymes are encoded by different genes (Schneider and Kennedy, 1967).

In this communication we describe the first isolation of cDNA clones encoding human ASM. Notably, two different groups of cDNA clones encoding ASM (from two different cDNA libraries) were identified by Southern hybridization analyses with different oligonucleotide mixtures. Furthermore, sequencing a representative of each group demonstrated that alternative splicing resulted in the occurrence of two ASM transcripts, presumably due to the presence of an intronic 3′ cryptic splice site.

©IRL Press

## Results

### Purification of human ASM

Human ASM was purified ~30 000-fold from 1000 l of urine by the method of Quintern *et al.* (1987). In the final preparation only the 72 kd monomeric form of ASM was detected by Coomassie blue staining (Figure 1A). The purified ASM was treated with glycopeptidase F and the mol. wt of the denatured, deglycosylated monomeric enzyme was ~61 kd (Figure 1B). From these results it was estimated that the purified urinary form of human ASM is composed of ~550 amino acid residues. This finding predicted a cDNA coding sequence of at least 1650 bp.

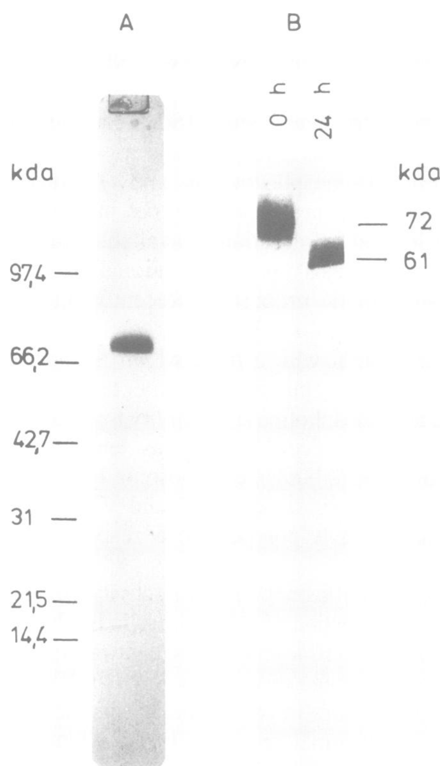


**Fig. 1.** SDS–PAGE of native and deglycosylated human ASM. (**A**) Purified urinary ASM (~2 μg) electrophoresed in a SDS–polyacrylamide gradient gel (3–23%) under non-reducing conditions. For the deglycosylation studies (**B**), ~0.5 μg of Endoglycosidase F was added to 2 μg of human ASM and incubated at 37°C for 0–24 h. The reaction mixture was then electrophoresed in a SDS–polyacrylamide gradient gel (3–23%).

### Amino acid sequencing and oligonucleotide synthesis

Since the N terminus of ASM was blocked, the purified urinary enzyme was digested with trypsin and the tryptic peptides were purified by HPLC. Twelve peptides were microsequenced (~20% of the purified enzyme) and four were chosen for the construction of corresponding oligonucleotide mixtures based on regions of minimal codon redundancy. No similarities were found with the ASM sequences and other protein sequences in the NBRF protein database. Table I shows the four tryptic peptides and the corresponding synthetic oligonucleotide mixtures that were constructed for cDNA library screening.

### Isolation of cDNA clones encoding ASM

Using oligonucleotide mixture 1 as a probe, $3 \times 10^6$ independent recombinants from a human fibroblast pcD library (Okayama and Berg, 1983) were screened. The 92 putative positive clones were isolated, purified and characterized by restriction enzyme and Southern analyses with oligonucleotide mixtures 2–4. Based on these results, two groups of clones were identified. Group I (80 clones with inserts ranging from ~1.2 to 1.6 kb) hybridized with oligonucleotide mixtures 1–4 probes, while group II (12 clones with inserts ranging from ~1.2 to 1.4 kb) hybridized with oligonucleotide mixtures 1–3. Intriguingly, some of the group II clones were larger than the group I clones, despite the fact that they did not hybridize with oligonucleotide mixture 4. Subsequently a human placental cDNA library was screened with a radiolabeled 450 bp cDNA fragment obtained from a group I insert. A total of 21 putative positive clones were purified and analyzed, resulting

**Table I.** Tryptic peptide sequences of human acid sphingomyelinase and corresponding oligonucleotide mixtures

| Tryptic peptide/Oligonucleotide | Sequence |
|---|---|
| T-1 | -S - W - S - W - N - Y - Y - R - |
| Oligo 1 (20mer; 256 mix) | TGG ACN TGG AAT TAT TAT AG<br>TG C C CC |
| T-2 | -A - W - E - P - W - L - P - A - E - L - R- |
| Oligo 2A[a] (17mer; 64 mix) | CGN ACC CTT GGN ACC AA<br>C G |
| Oligo 2B[a] (17mer; 64 mix) | ACC AAT GGN CGN CTT CG<br>C C |
| Oligo 2C[a] (17mer; 128 mix) | ACC GAN GGN CGN CTT CG<br>C |
| T-3 | -(X)- L - Y - E - A - M - A - K - |
| Oligo 3 (17mer; 64 mix) | TAT GAA GCN ATG GCN AA<br>C G |
| T-4 | -I - G - G - F - Y - A - L - S - P - Y - P - G - L - R - |
| Oligo 4 (17mer; 192 mix) | ATT GGN GGN TTT TAT GC<br>C A C C |

[a]Probes constructed as complement strand.

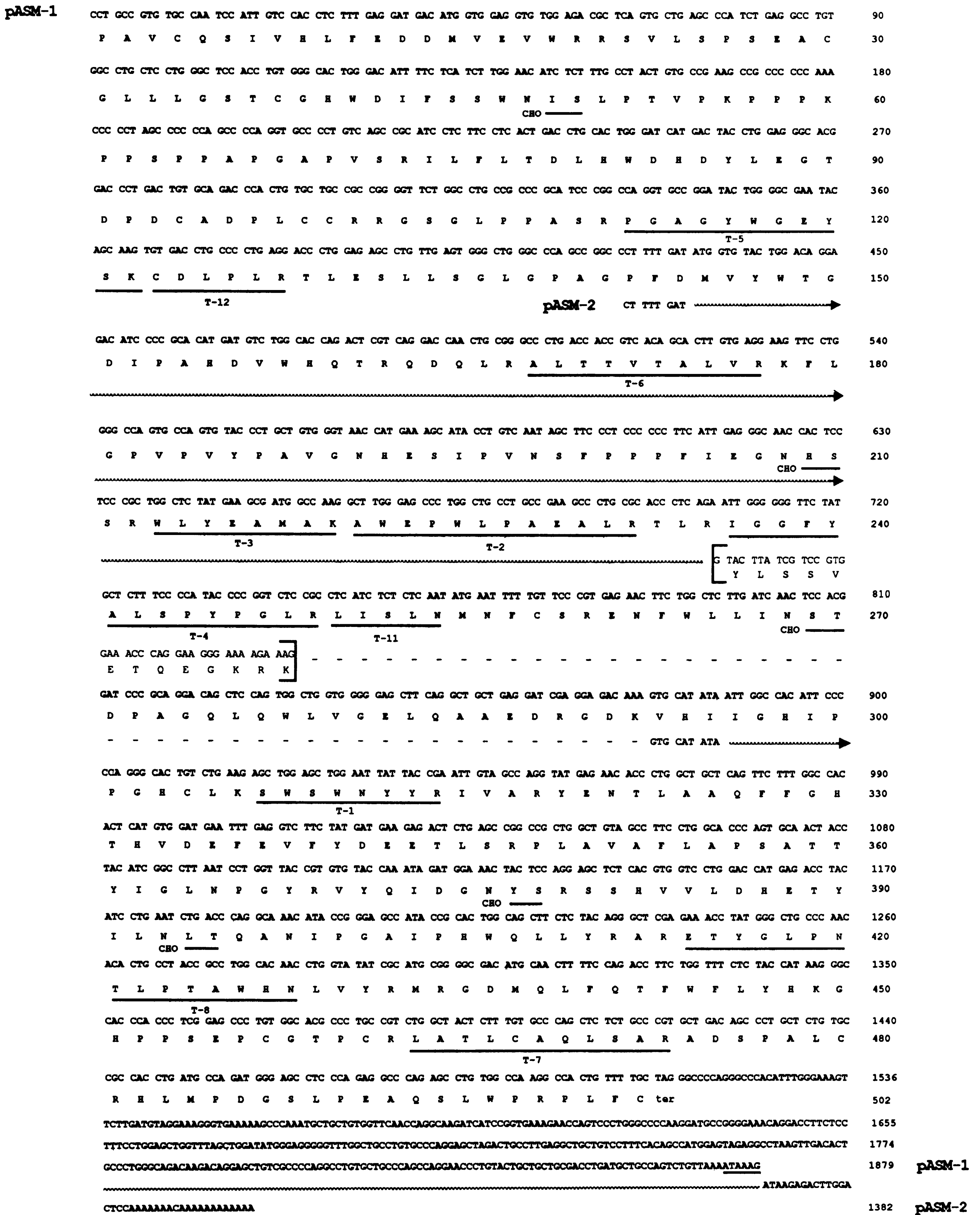


**Fig. 2.** Nucleotide and amino acid sequences of pASM-1 and pASM-2. The pASM-1 and pASM-2 inserts were 1879 and 1392 nucleotides (nt) long respectively. The wavey line indicates sequences that were identical in pASM-1 and pASM-2 from pASM-1 nt 424; the wavey line is not shown for pASM-1 nt 901 – 1773 although the two sequences were identical. Residues which were deleted in pASM-2 are indicated by dashes (–). Colinear tryptic peptides are underlined.

in the identification of 18 additional group I and three additional group II inserts.

Based on these findings, representative clones from group I (pASM-1) and group II (pASM-2) were selected for DNA sequencing by the dideoxy chain termination method (Figure 2). The sequence of clone pASM-1 (1879 bp), obtained from the placental cDNA library, was colinear with 96 microsequenced amino acid residues from human ASM, demonstrating its authenticity. There were three discrepancies between the microsequenced amino acid sequences from the tryptic peptides and those predicted from the cDNA sequence (residues 123, 252 and 467). No poly(A)$^+$ tail was found in pASM-1, although a consensus polyadenylation signal (AATAAA) was present at the 3′ terminus of the insert. ASM-2 (1382 bp) was colinear with 78 microsequenced amino acid residues and had the consensus polyadenylation signal and a poly(A)$^+$ tail (Figure 2). The polyadenylation consensus sequence was 27 bp from the poly(A) tract in pASM-2. Of particular interest, pASM-2 did not have an internal 172 bp sequence encoding 57 ASM amino acids which were present in pASM-1; this deleted sequence encoded the peptide corresponding to oligonucleotide mixture 4 (see Figure 2; T-4). In place of this sequence, pASM-2 contained an in-frame 40 bp sequence encoding 13 amino acids which were not present in pASM-1.

## Discussion

The recent purification of human ASM permitted the isolation of tryptic peptides for microsequencing and the synthesis of oligonucleotide mixtures based on regions of low codon redundancy (Table I). Using the oligonucleotide mixtures as probes, a total of 92 putative positive cDNA clones for human ASM were isolated and purified from a pcD human fibroblast library. Southern hybridization analyses revealed two groups of cDNA clones. Group I clones had inserts that hybridized with all four oligonucleotide mixtures, whereas the inserts of group II clones only hybridized with oligonucleotide mixtures 1–3, despite the fact that some of the group II inserts were larger than those found in group I. The nature of this apparent discrepancy was revealed by sequencing the longest cDNA insert from each group. pASM-1 (group 1) had an internal 172 bp sequence encoding 57 amino acids that was not present in pASM-2. This sequence contained the tryptic peptide used to construct oligonucleotide mixture 4 (Table I; T-4). In place of this sequence, pASM-2 had a 40 bp in-frame sequence that encoded 13 amino acid residues which were not present in pASM-1.

The finding of two authentic pASM cDNA clones that differed by the presence of two unique internal nucleotide sequences provided evidence for alternative splicing of the pASM transcript. The possibility that these results were caused by cloning artifacts was excluded by hybridization studies and sequencing of additional cDNA clones from human fibroblast and placental libraries. Both libraries contained group I and II ASM cDNA clones. Presumably, group I clones corresponded to the normally processed transcript that encodes human ASM, since they contained peptide T-4 and represented ~90% of the clones identified in the combined fibroblast and placental libraries. The fact that pASM-2 was missing 172 nucleotides but contained 40 in-frame nucleotides suggests that ~10% of the ASM transcripts are alternatively spliced. The transcript containing the unique 40 bp may result from alternative splicing of an exonic sequence, or is perhaps due to the presence of an intronic sequence that can serve as a cryptic 3′ donor splice site. Analysis of the unique 40 bp sequence in pASM-2 is consistent with either possibility. Note that the first (GT) and the last (AG) two nucleotides are the 5′ donor and 3′ acceptor consensus dinucleotides for intron splicing. The fact that the putative splicing event deletes 172 bp which are replaced by 40 bp (encoding 13 amino acids) without altering the reading frame suggests that both transcripts may encode functional proteins.

It is tempting to speculate on the possible significance of two functional human ASM transcripts. Since ASM has both sphingomyelinase and phospholipase C activities, it is possible that the two transcripts produce enzymes that recognize and preferentially bind different substrates. Alternatively, this region of the enzyme may be important for subcellular targeting since the sequences which are deleted in pASM-2 contain a potential N-glycosylation site. While ASM is a lysosomal hydrolase with an acidic pH optimum, the NSM activity has been localized to the plasma membrane. In fact, the alternatively spliced transcripts may result in an enzyme with different subcellular localization, pH optima and substrate specificity. The presence of two human transcripts also may provide insights into the clinical heterogeneity of NP disease. A mutation in the sequence which is deleted in pASM-2 would only affect the expression of the group I transcripts, leaving ~10% of the transcripts unaffected. If the group II transcripts express a functional protein this would result in residual ASM activity, consistent with the less severe NP-B phenotype. Mutations in common sequences would affect both transcripts and could result in little, if any, functional enzyme and the severe NP-A phenotype. On the other hand, the alternatively spliced transcript may not be functional. In fact, alternating splicing of the human β-glucuronidase transcripts supports this possibility (Oshima *et al.*, 1987). Of the two β-glucuronidase cDNAs isolated, one had a 153 bp deletion (corresponding to an entire exon) which expressed an enzyme protein that was not catalytically active. Clearly, transient expression of full-length ASM group I and II cDNAs will determine the functional integrity of the ASM transcripts.

Finally, since the nucleotide sequence of pASM-1 and pASM-2 are identical, except for the spliced region, it is unlikely that the two groups of clones represent transcripts of different ASM genes or pseudogenes. However, a complete understanding of the molecular nature of the putative alternative splicing event will require additional studies, including S1 nuclease analysis of the ASM RNA in different tissues and the isolation and sequencing of the implicated genomic region(s). Furthermore, the availability of cDNA clones encoding human ASM will permit studies of the structure and expression of human ASM, characterization of the genomic organization and regional chromosomal location of the ASM gene, as well as studies of the molecular nature of the mutations that cause the clinically distinct NP subtypes.

## Materials and methods

### *Materials*

Restriction endonucleases, polynucleotide kinase, T4 DNA ligase, glycopeptidase F and G-25 spin columns were from Boehringer Mannheim. Radioisotopes were from the Amersham Corporation. Nitrocellulose (type HATF)

and nylon membranes were from Millipore. M13 and Bluescript vectors were from New England Biolabs and Stratagene respectively. Reagents and columns for the synthesis of the oligonucleotide probes and sequencing primers were from Applied Biosystems.

#### *Purification of human ASM*

Human ASM was purified from 1000 l of urine according to the procedure of Quintern *et al.* (1987) with the following modifications. The dialyzed urine concentrate was loaded onto the octyl-Sepharose column without performing an ammonium sulfate precipitation as the first purification step. Fractions of the Blue-Sepharose eluate that contained ASM with the highest specific activity, were pooled and dialyzed against 40 mM Tris−HCl, pH 7.6, containing 0.1% NP-40 and loaded onto a DEAE−cellulose column equilibrated with the same buffer. Under these conditions, ASM did not bind to the column.

#### *Deglycosylation studies*

The purified enzyme was reduced with dithiothreitol and alkylated with iodoacetamide as previously described (Hempel *et al.*, 1984). Glycopeptidase F (0.5 mg) was added to 2 μg of the alkylated enzyme and incubated at 37°C for 0−24 h in 50 mM Tris−HCl, pH 7.5, containing 0.1% Nonidet P-40. Enzyme preparations were then analyzed by electrophoresis in SDS−polyacrylamide gradient gels (3−23%).

#### *Amino acid microsequencing*

Highly enriched ASM (~0.5 mg) was reduced with dithiothreitol, alkylated with iodoacetamide and hydrolyzed with trypsin (trypsin:protein ratio of 1:100). The tryptic peptides were separated on an HIBAR Lichrospher column (500 CH-8; 10 μm) with a 0−70% acetonitrile gradient in 0.05% trifluroacetic acid, and the single peptide peaks were then rechromatographed on the same column with a 0−70% acetonitrile gradient in 25 mM ammonium acetate buffer, pH 6.0. Only peptides which were eluted with high yields were sequenced by automated Edman degradation using an Applied Biosystems gas-phase amino acid sequencer with an on-line PTH analyzer (Beyreuther *et al.*, 1983; Stone and Williams, 1986).

#### *Synthesis of oligonucleotide mixtures*

Oligonucleotide mixtures corresponding to amino acid sequences of minimal codon redundancy were synthesized on an Applied Biosystems DNA Synthesizer using phosphoramidite chemistry (Itakura *et al.*, 1984; Caruthers, 1985). After synthesis, the mixtures were analyzed on 20% polyacrylamide−8 M urea gels prior to use and, if necessary, purified by isolating the band and eluting the DNA overnight. 5′-End labeling reactions were carried out using [$\gamma$-$^{32}$P]ATP (5000 Ci/mmol) using T4 polynucleotide kinase (Maniatis *et al.*, 1982).

#### *cDNA library screening*

The pcD human fibroblast cDNA library (Okayama and Berg, 1983) was plated at a density of ~$10^4$ colonies/150 $cm^2$ Petri dish and screened by standard methods (Maniatis *et al.*, 1982). Prehybridization (6 × SSPE, 10 × Denhart's solution, 0.5% SDS) and hybridization (6 × SSPE, 5 × Denhart's solution, 0.5% SDS containing 50 pmol $\gamma$-$^{32}$P-labeled oligonucleotide) were performed in pouches with 10 filters/bag at 6°C below the lowest $T_m$ for the oligonucleotide mixture. Filters were washed at the lowest $T_m$ for each oligonucleotide mixture in 6 × SSC, 0.1% SDS for 3−6 h.

#### *DNA sequencing and computer analyses*

cDNA inserts (and/or restriction fragments) from putative positive clones were subcloned into M13 (Messing *et al.*, 1981) and/or Bluescript (Stratagene, Inc.) vectors and sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977). Sequencing primers were synthesized on an Applied Biosystems DNA Synthesizer as described above. Computer analyses were performed using the University of Wisconsin Genetics Computer Group DNA Sequence Analysis Software.

## Acknowledgements

The authors wish to thank Dr H.Okayama, National Institutes of Health, for the human fibroblast pcD library, Dr Ken Williams at the Yale University School of Medicine and Dr K.Beyreuther at DKFZ Heidelberg for the amino acid microsequencing, and Ms Christine Jackson for valuable assistance. This work was supported by a grant (5 RO1 DK34045) from the National Institutes of Health, a grant from the Deutsche Forschungsgemeinschaft and a grant from the Fonds der Chemischen Industrie. E.H.S. is the recipient of a Basil O'Conner Starter Scholar Research Award (5-640) and a Charles H.Revson Foundation Fellowship in Biomedical Sciences.

## References

Brady,R.O. (1983) In Stanbury,J.B., Wyngaarden,J.B., Fredrickson,D.S., Goldstein,J.L. and Brown,M.S. (eds), *The Metabolic Basis of Inherited Diseases*, 7th edn. McGraw-Hill, New York, pp. 831−884.

Besley,G.T.N., Hoogeboom,A.J.M., Moogeveen,A., Kleijger,W.J. and Galjaard,H. (1980) *Hum. Genet.*, **54**, 409−412.

Beyreuther,K., Bieseler,B., Bovens,J., Dildrop,R., Neifer,K., Stüber,K., Zaiss,S. and Ehring,R. (1983) *Modern Methods in Protein Chemistry.* Walter de Gruyter, Berlin, pp. 303−325.

Caruthers,M.H. (1985) *Science*, **230**, 281−285.

Hempel,J., von Bahr-Lindström,H. and Jörnvall,H. (1984) *Eur. J. Biochem.*, **141**, 21−35.

Itakura,K., Rossi,J.J. and Wallace,R.B. (1984) *Annu. Rev. Biochem.*, **53**, 323−356.

Kitagawa,T. (1987) *Jpn. J. Human Genet.*, **32**, 55−69.

Konrad,R. and Wilson,D. (1987) *Cytogenet. Cell Genet.*, **46**, 641.

Levade,T., Salvayre,R. and Douste-Blazy,L. (1986) *J. Clin. Chem. Clin. Biochem.*, **24**, 205−220.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: a Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.*, **9**, 309−321.

Okayama,H. and Berg,P. (1983) *Mol. Cell. Biol.*, **3**, 280−289.

Oshima,A., Kyle,J.W., Miller,R.D., Hoffman,J.W., Powell,P.P., Grubb,J.H., Sly,W., Tropak,M., Guise,K.S. and Gravel,R.A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 685−689.

Quintern,L.E., Weitz,G., Nehrkorn,H., Tager,J.M., Schram,A.W. and Sandhoff,K. (1987) *Biochim. Biophys. Acta*, **922**, 323−336.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463−5467.

Schneider,P.B. and Kennedy,E. (1967) *J. Lipid Res.*, **8**, 202−209.

Stone,K. and Williams,K. (1986) *J. Chromatogr.*, **359**, 203−212.

*Received on March 18, 1989; revised on May 22, 1989*