Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 3660–3664, April 1993
Applied Biological Sciences

Cloning and expression of heparinase I gene from *Flavobacterium heparinum*

(heparin/PCR/extracellular matrix)

RAM SASISEKHARAN*†, MARK BULMER‡, KELLEY W. MOREMEN§¶, CHARLES L. COONEY‡, AND ROBERT LANGER*‡

†Division of Medical Sciences, Harvard Medical School, Boston, MA 02115; and §Center for Cancer Research, ‡Department of Chemical Engineering, and *Harvard-Massachusetts Institute of Technology Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Robert Langer, December 29, 1992

ABSTRACT **Heparinases, enzymes that cleave heparin and heparan sulfate, are implicated in physiological and pathological functions ranging from wound healing to tumor metastasis and are useful in deheparinization therapies. We report the cloning of the heparinase I (EC 4.2.2.7) gene from *Flavobacterium heparinum* using PCR. Two degenerate oligonucleotides, based on the amino acid sequences derived from tryptic peptides of purified heparinase, were used to generate a 600-bp probe by PCR amplification using *Flavobacterium* genomic DNA as the template. This probe was used to screen a *Flavobacterium* genomic DNA library in pUC18. The open reading frame of heparinase I is 1152 bp in length, encoding a precursor protein of 43.8 kDa. Eleven of the tryptic peptides (≈35% of the total amino acids) mapped onto the open reading frame. The amino acid sequence reveals a consensus heparin binding domain and a 21-residue leader peptide with a characteristic Ala-(Xaa)-Ala cleavage site. Recombinant heparinase was expressed in *Escherichia coli* as a soluble protein, using the T7 polymerase pET expression system. The recombinant heparinase cleavage of heparin was identical to that of native heparinase.**

Heparin-like molecules, mammalian complex carbohydrates, are believed to be present on virtually all cell surfaces and are associated with membrane proteins or proteoglycans (1). Heparin not only provides architecture and hydration to the extracellular matrix but also interacts with an array of proteins, such as growth factors, present in the extracellular matrix (2, 3). Heparin acts as a mesh, in sequestering and protecting cytokines from proteolytic degradation and in facilitating their binding to receptors (4). Chemically, heparin and the related heparan sulfate are characterized by a disaccharide repeating unit of uronic acid (L-iduronic acid or D-glucuronic acid) and hexosamine connected through a 1,4 linkage. Heparin is highly heterogeneous in its composition due to the varying location of sulfate substituents (5). The best-characterized aspect of heparin is its anticoagulant property; it is widely used as an anticoagulant in surgery (6).

Heparinases degrade heparin and heparan sulfate; they are implicated also in the release of heparin-bound growth factors from the extracellular matrix (7). These enzymes are important tools in understanding heparin structure and function (8). Heparinases are either lyases or hydrolases. Relatively few sources of these enzymes are known (9). *Flavobacterium heparinum*, a Gram-negative, nonpathogenic soil organism, produces enzymes that degrade and modify heparin (10). Interestingly, *F. heparinum* can use heparin as its sole carbon source. It produces three heparinases—heparinase I (EC 4.2.2.7), a 42.5-kDa enzyme that acts at the hexosamine–iduronic acid linkage of the saccharide; heparinase II (no EC number), an 84-kDa enzyme that acts at the hexosamine–uronic acid linkage, not discriminating between the two isoforms of the uronic acid; and heparinase III (EC 4.2.2.8), a 70-kDa enzyme that acts at the hexosamine–glucuronic acid linkage (11).

Heparinase I expression in *F. heparinum* is induced by heparin in the initial stages of heparin catabolism (12). It depolymerizes heparin in a random endolytic fashion into di-, tetra-, and hexasaccharides, with a reducing and unsaturated 4,5 end group (13). It is isolated from the periplasm of the *F. heparinum* as a monomeric protein with an apparent molecular mass of 43 kDa and a pI of 8.5 (14, 15). Initial studies established that complete enzymatic degradation of heparin by heparinase I was sufficient to eliminate the anticoagulant properties of heparin in surgery and this led to the proposal of a novel postsurgical therapy for deheparinization (16).

We report the cloning and expression, in *Escherichia coli*, of heparinase I from *F. heparinum*.‖ To our knowledge, no other gene has been cloned from *F. heparinum*, and this is the first report of the cloning of a heparin-degrading enzyme.

MATERIALS AND METHODS

Materials. Heparin, from porcine intestinal mucosa, 157 USP units/mg, was from Hepar, Franklin, OH. It was prepared at concentrations of 25 mg/ml or 2 mg/ml in 5 mM calcium acetate and in 100 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS) buffer (pH 7.0). *E. coli* BL(DE3) host was from Novogen, Madison, WI. *E. coli* DH5α was from GIBCO/BRL/Life Technologies. Biochemical and molecular biology reagents and their sources are listed in the appropriate sections. Native heparinase was purified from *F. heparinum* (15, 16). Urea, dithiothreitol, iodoacetamide, trifluoroacetic acid (TFA), and acetonitrile were from Allied Chemicals. Trypsin was from Pierce. The other chemicals were from Mallinckrodt.

Heparinase Purification and Characterization. Heparinase was purified (15, 17, 18) and an additional purification step was carried out by high-pressure liquid chromatography (HPLC) (in a HP 1090 from Hewlett–Packard, with diode array detection) using a Vydac C_{18} reverse-phase column, with a gradient of 0–80% acetonitrile in 0.1% TFA for 60 min. Protein was monitored at 210 and 277 nm. Heparinase I appeared as a doublet and was further separated using a shallow gradient of 0–40% acetonitrile for 20 min, followed by 40–80% acetonitrile for 40 min. The separated heparinase

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: r-heparinase, recombinant heparinase; ORF, open reading frame; TFA, trifluoroacetic acid.

¶Present address: Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602.

‖The sequence reported in this paper has been deposited in the GenBank data base (accession no. L12534).

peaks were collected in a Microfuge tube and lyophilized (VirTis Freeze Mobil model 12, VirTis). Protein concentration was determined by use of the micro BCA (bicinchoninic acid) reagent (Pierce) relative to a bovine serum albumin standard. The major isoform was used in the tryptic digest. Mass spectrometry was performed on the major heparinase isoform preparations to determine the purity, homogeneity, and molecular mass of heparinase I. About 2 μg of heparinase was mixed with 1 μl of sinapinic acid (10 mg/ml; in 80% acetonitrile/0.1% TFA in water), in equal vol/vol ratio, and then analyzed using laser desorption mass spectrometry (Laser MAT, Finnigan-MAT, San Jose, CA) (19).

Tryptic Digest and Protein Sequence Analyses. One nanomole (≈40 μg) of the purified enzyme (the major isoform indicated above) was denatured in 50 μl of 8 M urea/0.4 M ammonium carbonate and reduced with 5 mM dithiothreitol at 50°C, cooled to room temperature, and alkylated with 10 mM iodoacetamide for 15 min in the dark. The reaction was quenched with water by bringing the total reaction volume to 200 μl. To the above reaction, 4% (wt/wt) trypsin was added and the digestion was carried out at 37°C for 24 hr. The proteolytic reaction was terminated by heating the sample at 65°C for 2 min. The digest was separated using a gradient reverse-phase HPLC (0–80% acetonitrile in 0.1% TFA for 120 min). Tryptic peptides were monitored at 210 and 277 nm and collected in Microfuge tubes. Based on the homogeneity of the peptide peaks, eight different peaks were sequenced using an Applied Biosystems sequencer model 477, with an on-line model 120 phenylthiohydantoin amino acid analyzer (Biopolymers Laboratory, Center for Cancer Research, Massachusetts Institute of Technology). Native undigested heparinase also was analyzed to determine the N-terminal sequence.

Genomic DNA Isolation, Plasmid Library Preparation, and Southern Blotting. The *F. heparinum* genomic DNA was isolated by the A.S.A.P. kit (Boehringer Mannheim) with the following modifications. The DNA was desalted and concentrated using a Centricon P-30 (Amicon) to a final volume of 100 μl in water. For 10^9 cells, 105–115 μg of DNA typically was obtained. The *F. heparinum* genomic DNA plasmid library was a generous gift from A. J. Sinskey (Massachusetts Institute of Technology). A *F. heparinum* genomic DNA library was prepared by sonicating 150 μg of genomic DNA, followed by the addition of *Eco*RI linkers and ligation into pUC18. The pUC18 genomic library was transformed in the *E. coli* host DH5α (20). Genomic DNA (6 μg) was prepared for Southern blotting by digestion with *Eco*RI, *Bam*HI, and *Hin*dIII, individually or in combination, for 2 hr, and separated on a 0.8% agarose gel for 16 hr (60 V). The gel was transferred onto a nylon membrane (Hybond-ECL, Amersham) by capillary action (20). The probe was prepared using the ECL labeling kit as per the manufacturer's recommendations (Amersham). The membrane was blocked, hybridized, probed using the 600-bp PCR product, and detected using the ECL kit as per the manufacturer's recommendations (Amersham).

Amplification of the PCR Product. DNA amplification using heparinase primers was carried out in a 25-μl reaction volume containing 3 μg of the genomic DNA as the template, 50 mM KCl, 10 mM Tris·HCl (pH 8.3), 1.5 mM $MgCl_2$, and 0.01% gelatin with the four dNTPs, at 200 μM, 0.5 μM primer, 2.5 units of the *Taq* polymerase (Cetus), and 25 μl of mineral oil. The samples were placed in an automated heating block (DNA thermal cycler, Perkin–Elmer) programed for step cycles of 92°C (2 min), 50°C (1 min), and 72°C (3 min). This cycle was repeated 35 times and the final cycle had a 10-min extension at 72°C. The control reaction was provided by the Cetus kit. The PCR products were analyzed on a 0.8% agarose gel. The PCR product was isolated on a low-melting agarose gel using the GeneClean procedure (Bio 101, La Jolla, CA).

Screening of the *F. heparinum* pUC18 Genomic Library. A *F. heparinum* genomic DNA plasmid library (1500 colonies) was screened by colony hybridization, with a radiolabeled 600-bp heparinase PCR product as the probe (20). Approximately 100 colonies were grown directly on nitrocellulose filters on LB ampicillin plates and then replicated on additional nitrocellulose filters and grown on fresh LB ampicillin plates overnight. The heparinase PCR product was isolated as described above and labeled using the random hexanucleotide kit (IBI) and [α-^{32}P]dCTP (3000 Ci/mmol; 1 Ci = 37 GBq) (New England Nuclear). Following labeling of the probe, free nucleotides were removed with a Sephadex G-50 column to a final volume of 100 μl in water (Nick Column, Pharmacia). About 10^6 cpm of the probe per ml was used for screening.

DNA Sequencing. DNA sequencing was performed in phage M13 employing the dideoxyadenosine 5′-[α-[^{35}S]thio]-triphosphate (NEN) and Sequenase (United States Biochemical) as described by the manufacturer. The sequence data were obtained using successive nested deletions in M13 with T4 DNA polymerase, as per Cyclone I Biosystems (IBI), or sequenced using synthetic oligonucleotide primers designed for different regions of the heparinase gene in both orientations. The DNA sequence was analyzed with MACVECTOR software on Macintosh Fx (IBI).

Recombinant Heparinase (r-Heparinase). r-Heparinase was expressed using the pET system (21), where expression is driven by bacteriophage T7 polymerase. The host *E. coli* strain BL21(DE3) contains a chromosomal copy of the T7 polymerase gene under the control of *lac*UV5. The expression is induced by isopropyl β-D-thiogalactoside. Two constructs were designed to express heparinase in the pET system. The first construct included the native heparinase leader sequence. The second construct started with a sequence that read Met, Gln_{22}, Gln_{23}, Lys_{24}, Lys_{25}, Ser_{26} . . . . The Met residue was added before the Gln_{22} to introduce a start codon. *Nde* I and *Bam*HI restriction sites were appended onto the 5′ ends of the N- and C-terminal primers, respectively. *F. heparinum* genomic DNA was used as a template for 10 scaled-up PCR cycles (see Amplification of the PCR Product) to generate the modified heparinase I product. The PCR product was isolated as discussed above and concentrated using a Centricon P-100 (Amicon) at 5000 × *g* for 20 min. The amplification product was treated with Klenow fragment of DNA polymerase I (New England Biolabs) for 15 min and with T4 DNA polymerase (New England Biolabs) for 10 min. The PCR product was isolated on an agarose gel as described above. The blunt PCR fragment was concentrated using a Centricon P-100 and ligated with 20 ng of *Sma* I-digested pUC18 (Pharmacia) (19). The ligation mixture was then used to transform DH5α-competent cells (GIBCO/BRL/Life Technologies), as recommended by the manufacturer. Subcloned heparinase PCR fragments were excised from pUC18 by digestion with *Nde* I and *Bam*HI, gel purified, and then ligated to pET-3a plasmid (predigested at the *Nde* I and *Bam*HI sites and gel purified) using T4 DNA ligase (New England Biolabs). The ligation mixture then was used to transform DH5α-competent cells. The plasmid containing the heparinase I gene in pET-3a was isolated, purified, and used to transform the host cell BL21(DE3) (Novogen, Madison, WI).

Overnight BL21(DE3) cultures containing the heparinase I gene fragment in pET-3a were further diluted to an OD_{600} of 0.5–0.6 and then grown to an OD_{600} of 1.0, at 37°C, followed by induction with 1 mM isopropyl β-D-thiogalactoside. The induction was carried out for 2 hr, at 37°C; after 2 hr, 40 ml (10^8 cells per ml) of cells was centrifuged and washed twice with phosphate-buffered saline (pH 7.0). The cells were

reconstituted in a 5-ml solution of 50 mM Tris·HCl/2 mM EDTA, pH 8.0, and homogenized using a Polytron homogenizer (setting 7, in 30-sec bursts, for a total of 8 min). Nucleic acids were precipitated using 0.5% protamine sulfate and the samples were clarified by centrifugation at 5000 × *g* using a Beckman centrifuge (model no. GTKR) for 20 min. The samples were desalted and concentrated using a Centricon P-30. The final volume of the cell extract was about 100 μl. Then 5 μl of the concentrate was used in the enzyme activity assay. About 20 μl (in 2 ml) of the Centricon-concentrated r-heparinase extract was loaded onto a POROS HS/M (4.6 mm × 100 mm) cation-exchange column (PerSeptive Biosystems, Cambridge, MA). Each run was for 10 min using a salt gradient of 0–1 M NaCl (in 5 min) in 5 mM Tris·HCl (pH 7.0), monitored at 210 nm. As a control, 2 μg of the *F. heparinum* heparinase was run through the column, to determine the retention time. Peaks in the r-heparinase run, corresponding to the same retention time as the control heparinase run, were isolated and then desalted and concentrated using a Centricon P-30. Five microliters each of concentrated material was used in the enzyme assay and in SDS/PAGE analysis.

Gel Electrophoresis. SDS/PAGE of heparinase (*F. heparinum* and the r-heparinase) was carried out on a 12% polyacrylamide gel and separated using a Mini Protean II electrophoresis apparatus (120 V for 90 min) (Bio-Rad) (22). Molecular weight standards were obtained from GIBCO/BRL/Life Technologies. Proteins were visualized with a 0.1% Coomassie brilliant blue R-250 solution followed by destaining with a 40% (vol/vol) methanol/10% (vol/vol) acetic acid aqueous solution.

Heparinase Assay. Ten microliters of enzyme solution was added to 2 mg or 25 mg of heparin per ml in 5 mM calcium acetate/100 mM MOPS buffer (pH 7.0) in a total volume of 900 ml. The reaction mixture was incubated at 37°C. At various time intervals (for a total of 16 hr), aliquots of 50 μl were withdrawn in duplicate, to determine the increase in absorbance at 232 nm, as described by Bernstein *et al.* (18). Heparin (2 mg/ml) was degraded by the *F. heparinum* heparinase and r-heparinase in the 5 mM calcium acetate/100 mM MOPS buffer, pH 7.0, for 18 hr. The reaction was stopped by injecting the solution into a POROS Q/M (4.6 mm × 100 mm) anion-exchange column (PerSeptive Biosystems) connected to a BioCAD system (PerSeptive Biosystems). Each run was for 10 min using a salt gradient of 0–1 M NaCl (in 5 min) in 5 mM Tris·HCl (pH 7.0), monitored at 232 nm. The control heparin digest was carried out with 0.2 μg of *F. heparinum* heparinase. Heparin digest using r-heparinase was with 5 μl of the *E. coli* crude extract, having a total protein concentration of about 4.2 μg/μl. Anion-exchange HPLC resolves the oligosaccharide products (di-, tetra-, hexa-, and higher saccharides) produced upon heparin degradation by heparinase (13).

RESULTS

Heparinase Purification and Homogeneity. The reverse-phase HPLC purification of heparinase resulted in a doublet. This doublet was further separated, using a shallow gradient, into a major and a minor peak corresponding to two heparinase isoforms (Fig. 1*A*). The minor isoform eluted as a hydrophilic protein compared to the major isoform. The major isoform was isolated and used in this work. Laser desorption mass spectrometry of the major heparinase isoform indicated that the protein was a homogeneous preparation (data not shown). The average molecular mass of heparinase I estimated by mass spectrometry is 42.575 ± 0.120 kDa.

Cloning Strategy, PCR Amplification, and Screening. The HPLC-purified heparinase (Fig. 1*A*) was reduced, alkylated, and digested with trypsin, and rechromatographed by reverse-phase HPLC, to obtain ≈60 peptide peaks (Fig. 1*B*). These peaks were monitored at 210 and 277 nm, collected, and sequenced. Of the six peaks initially sequenced (Table 1), three were chosen for primer design (Table 2) (23). The PCR product of the combination of the primers td43 and td33 was about 150 bp in length. The combination of td4 and td33 primers was about 600 bp. Using the PCR product of td4 and td33 as a template, and td43 and td4 as primers, the predicted 150-bp product was obtained. This confirmed that td43 was internal to td4 and td33. Based on the SDS/PAGE mobility of purified heparinase I, the molecular mass was estimated as 43 kDa. This would mean that the 600-bp PCR product represents about 51% of the approximated total of 1170 bp for the heparinase I gene. The 600-bp probe was chosen for screening of the pUC18 library by high stringency colony hybridization. Two positive clones were identified and isolated by three rounds of colony screening.

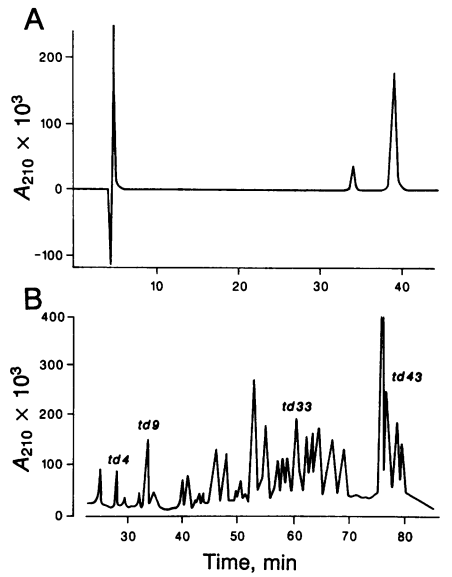

FIG. 1. Reverse-phase HPLC profile of purified heparinase I and tryptic digest of heparinase I. (*A*) Heparinase I appears as a doublet and the separation of the two isoforms of heparinase is shown. The major isoform was isolated and digested by trypsin. (*B*) Heparinase digested with trypsin, resolved on a reverse-phase HPLC column, and monitored at 210 nm. The peaks isolated and sequenced for cloning are td9, td33, and td43.

Southern Blotting, Restriction Mapping, and Sequencing of the Heparinase Clone. Southern blots, using *F. heparinum*

Table 1. Amino acid sequences of the tryptic peptides of heparinase

Peptide	Amino acid sequence
td04	**(K,R)GICEQGSSR**
td09	**(K,R)TVYHYGK**
td09′	**(K,R)TSTIAYK**
td21	**(K,R)FGIYR**
td33	**(K,R)ADIVNQQEILIGRDD*GYYFK**
td39	**(K,R)ITYVAGKPNGNKVEQGGYPTLAF***
td43	**(K,R)MPFAQFPKDCWITFDVAID*TK**
td40	**(K,R)NLSGYSETAR**
tdm4	KNIAHDKVEKK
td72	KTLSIEEFLALYDR
tdLx	RSYTFSVYIPSSFPDNATTIFAQWHGAPSRTLVA TPEGEIK

All of the heparinase I tryptic peptides sequenced are listed. Boldface type indicates the peptides were sequenced before the gene was cloned. The sequence begins **(K,R)** because trypsin cuts at either lysine or arginine residues. *, Amino acids that could not be determined.

Table 2. Primer design for heparinase

Peptide	Amino acid sequence
td04	**KGICEQGSSR**
y1	5′-AAA GGI AT(T/C/A) TGZ GAW CAW GG-3′
y2	5′-CC ZTG ZTC WCA (T/G/A)AT ICC TTT-3′
td43	(K,R)**MPFAQFPKDC**WITFCVAID*TK
D	5′-ATG CCI TTZ GCI CAW TTZ CCI AAW GAZ TG-3′
E	3′-TAC GGI AAW CGI GTZ AAW GGI TTZ CTW AC-5′
td33	(K,R)AD**IVNQQEILIG**RNDD*GYYFKA
A	5′-ATI AAZ CAW GAW ATI ZTI AT(T/C/A) GG-3′
B	5′-CC IAT IAW IAT ZTC ZTG ZTG WTT IAC XAT-3′
C	5′-CC IAT IAW IAT ZTC ZTG ZTG WTT IAC YAT-3′
	W = A/G X = A/C
	Y = T/G Z = T/C

The three sets of primers used in the PCR for cloning of the heparinase I gene are indicated. For the primer here I is the nucleotide deoxyinosine. Amino acids in boldface type represent the residues chosen for the primer design. Two different sets were constructed for tryptic peptide 33 to reduce the inosine substitution at the 3′ end of the primer. *, Amino acids that could not be determined.

genomic DNA and the 600-bp PCR product as the probe, indicated that the heparinase I gene was contained in an ≈15-kb *Eco*RI fragment and in an ≈3.5-kb *Hin*dIII fragment of the genomic DNA. The two positive clones (pRS51hep and pRS43hep), identified from the above plasmid screening, were characterized by restriction mapping. One clone, pRS51hep, contained a 2.3-kb insert, with a 1.6-kb *Kpn* I–*Kpn* I fragment. This fragment tested positive as a template for generating the 600-bp PCR product. The *Kpn* I–*Kpn* I fragment was subcloned into M13 and sequenced. The sequence (Fig. 2) reveals a single, continuous open reading frame (ORF) of 1152 bp (384 amino acids) containing a leader sequence of about 21 amino acids. The PCR product spans bases 566–1216 (from the ATG start site) and corresponds to about 57% of the total gene. Initially, six different tryptic peptides mapped onto the ORF (Table 1). Subsequently, five other peptides that were sequenced also mapped onto the ORF, together constituting about 34% of the total of 384 amino acids. There are three cysteine residues, of which one was associated with the signal peptide. A consensus heparin binding domain was identified in the primary sequence (24). The signal peptide is typical of prokaryotic signal sequences, with a consensus Ala-(Xaa)-Ala site for cleavage (25).

r-Heparinase. r-Heparinase expression in the pET system, using the *F. heparinum* leader sequence, was unsuccessful. Interestingly, N-glycosidase F, a carbohydrate-degrading enzyme produced by *Flavobacterium meningosepticum*, was not successfully expressed in *E. coli* with its native leader sequence of about 45 amino acids (26). The construct devoid of the native leader sequence was successful in expressing active heparinase I. r-Heparinase production steadily increased until 2 hr after induction (Fig. 3*A*). As seen from the SDS/PAGE gel, the r-heparinase migrated faster than *F. heparinum* heparinase. r-Heparinase was expressed as a soluble protein in *E. coli* with an activity of ≈3 units/mg of *E. coli* crude extract. The crude extract, with as little as 4 μg of the total protein, exhibited heparinase activity. r-Heparinase degradation of heparin was identical to that of the purified *F. heparinum* heparinase, producing the di-, the three tetra-, and the hexasaccharides (Fig. 3*B*). Purification of r-heparinase by cation-exchange chromatography enriched the enzyme as seen from SDS/PAGE (data not shown). It was interesting to note that although many *E. coli* cellular proteins did not bind to the cation-exchange column, the r-heparinase activity was found to bind and elute with a retention time similar to that of the *F. heparinum* heparinase. The SDS/PAGE separation of the cation fraction exhibiting enzymatic activity showed the presence of the 42.5-kDa species for the *F. heparinum* heparinase and the *E. coli* extract. The specific activity of r-heparinase (1 unit/5.2 μg) was comparable to that of the *F. heparinum* heparinase (1 unit/4 μg) (18).

DISCUSSION

We report here the cloning of *F. heparinum* heparinase I gene. Heparinase was purified to homogeneity by reverse-phase HPLC and the isoforms were separated. The major heparinase isoform was digested by trypsin and resulted in about 60 peaks, consistent with the total of 54 Lys and Arg residues in *F. heparinum* heparinase. Three peptides were chosen for primer design: td4, td33, and td43. The 600-bp PCR product, generated by the combination of td4 and td33, was validated by reamplification using the primers td43 and td33, since td43 was contained 150 bp upstream of td33 within the 600-bp td4-td33 PCR product. A *F. heparinum* genomic DNA library in pUC18 was screened using the 600-bp probe. Two colonies were initially isolated, with the clone pRS51hep containing a 2.3-kb insert. A 1.6-kb *Kpn* I–*Kpn* I fragment

```
1    CCTTT TGGGA GCAAA GGCAG AACCA TCTCC GAACA AAGGC AGAAC CAGCC TGTAA ACAGA CAGCA ATTCA TCCGC TTTCA ACCAA AGTGA AAGCA TTTAA TACAA TACCA GAATG TCGCA  120

121  TTTCC CTTTC AGCGT ACTTT TTGGG TAAAT AACCA ATAAA AACTA AAGAC GG ATG AAA  AAA CAA ATT CTA TAT CTG ATT GTA CTT CAG CAA CTG TTC CTC TGT TCG GCT TAC  232
                                                                  Met Lys  Lys Gln Ile Leu Tyr Leu Ile Val Leu Gln Gln Leu Phe Leu Cys Ser Ala Tyr
233  GCC CAG CAA AAA AAA TCC GGT AAC ATC CCT TAC CGG GTA AAT GTG CAG GCC GAC AGT GCT AAG CAG AAG GCG ATT ATT GAC AAC AAA TGG GTG GCA GTA GGC ATC AAT  340
     Ala Gln Gln Lys Lys Ser Gly Asn Ile Pro Tyr Arg Val Asn Val Gln Ala Asp Ser Ala Lys Gln Lys Ala Ile Ile Asp Asn Lys Trp Val Ala Val Gly Ile Asn
341  AAA CCT TAT GCA TTA CAA TAT GAC GAT AAA CTG CGC TTT AAT GGA AAA CCA TCC TAT CGC TTT GAG CTT AAA GCC GAA GAC AAT TCG CTT GAA GGT TAT GCT GCA GGA  448
     Lys Pro Tyr Ala Leu Gln Tyr Asp Asp Lys Leu Arg Phe Asn Gly Lys Pro Ser Tyr Arg Phe Glu Leu Lys Ala Glu Asp Asn Ser Leu Glu Gly Tyr Ala Ala Gly
449  GAA ACA AAG GGC CGT ACA GAA TTG TCG TAC AGC TAT GCA ACC ACC AAT GAT TTT AAG AAA TTT CCC CCA AGC GTA TAC CAA AAT GCG CAA AAG CTA AAA ACC GTT TAT  556
     Glu Thr Lys Gly Arg Thr Glu Leu Ser Tyr Ser Tyr Ala Thr Thr Asn Asp Phe Lys Lys Phe Pro Pro Ser Val Tyr Gln Asn Ala Gln Lys Leu Lys Thr Val Tyr
557  CAT TAC GGC AAA GGG ATT TGT GAA CAG GGG AGC TCC CGC AGC TAT ACC TTT TCA GTG TAC ATA CCC TCC TCC TTC CCC GAC AAT GCG ACT ACT ATT TTT GCC CAA TGG  664
     His Tyr Gly Lys Gly Ile Cys Glu Gln Gly Ser Ser Arg Ser Tyr Thr Phe Ser Val Tyr Ile Pro Ser Ser Phe Pro Asp Asn Ala Thr Thr Ile Phe Ala Gln Trp
665  CAT GGT GCA CCC AGC AGA ACG CTT GTA GCT ACA CCA GAG GGA GAA ATT AAA ACA CTG AGC ATA GAA GAG TTT TTG GCC TTA TAC GAC CGC ATG ATC TTC AAA AAA AAT  772
     His Gly Ala Pro Ser Arg Thr Leu Val Ala Thr Pro Glu Gly Glu Ile Lys Thr Leu Ser Ile Glu Glu Phe Leu Ala Leu Tyr Asp Arg Met Ile Phe Lys Lys Asn
773  ATC GCC CAT GAT AAA GTT GAA AAA AAA GAT AAG GAC GGA AAA ATT ACT TAT GTA GCC GGA AAG CCA AAT CGC TGG AAG GTA GAA CAA GGT GGT TAT CCC ACG CTG GCC  880
     Ile Ala His Asp Lys Val Glu Lys Lys Asp Lys Asp Gly Lys Ile Thr Tyr Val Ala Gly Lys Pro Asn Gly Trp Lys Val Glu Gln Gly Gly Tyr Pro Thr Leu Ala
881  TTT GGT TTT TCT AAA GGG TAT TTT TAC ATC AAG GCA AAC TCC GAC CGG CAG TGG CTT ACC GAC AAA GCC GAC CGT AAC AAT GCC AAT CCC GAG AAT AGT GAA GTA ATG  988
     Phe Gly Phe Ser Lys Gly Tyr Phe Tyr Ile Lys Ala Asn Ser Asp Arg Gln Trp Leu Thr Asp Lys Ala Asp Arg Asn Asn Ala Asn Pro Glu Asn Ser Glu Val Met
989  AAG CCC TAT TCC TCG GAA TAC AAA ACT TCA ACC ATT GCC TAT AAA ATG CCC TTT GCC CAG TTC CCT AAA GAT TGC TGG ATT ACT TTT GAT GTC GCC ATA GAC TGG ACG 1096
     Lys Pro Tyr Ser Ser Glu Tyr Lys Thr Ser Thr Ile Ala Tyr Lys Met Pro Phe Ala Gln Phe Pro Lys Asp Cys Trp Ile Thr Phe Asp Val Ala Ile Asp Trp Thr
1097 AAA TAT GGA AAA GAG GCC AAT ACA ATT TTG AAA CCC GGT AAG CTG GAT GTG ATG ATG ACT TAT ACC AAG AAT AAG AAA CCA CAA AAA GCG CAT ATC GTA AAC CAG CAG 1204
     Lys Tyr Gly Lys Glu Ala Asn Thr Ile Leu Lys Pro Gly Lys Leu Asp Val Met Met Thr Tyr Thr Lys Asn Lys Lys Pro Gln Lys Ala His Ile Val Asn Gln Gln
1205 GAA ATC CTG ATC GGA CGT AAC GAT GAC GAT GGC TAT TAC TTC AAA TTT GGA ATT TAC AGG GTC GGT AAC AGC ACG GTC CCG GTT ACT TAT AAC CTG AGC GGG TAC AGC 1312
     Glu Ile Leu Ile Gly Arg Asn Asp Asp Asp Gly Tyr Tyr Phe Lys Phe Gly Ile Tyr Arg Val Gly Asn Ser Thr Val Pro Val Thr Tyr Asn Leu Ser Gly Tyr Ser
1313 GAA ACT GCC AGA TAG CAA AAGCC CTAAG CGCAT CCGAT AGGGC TTTTC TTATA TTTAC AATAA AATT 1379
     Glu Thr Ala Arg ***
```

FIG. 2. ORF encoding the gene for heparinase I from *F. heparinum*. The three primers used in the cloning are solidly underlined (——). The peptides corresponding to these regions, td4, td33, and td43, are underlined (----). The signal peptide is indicated by a dotted line (·····).

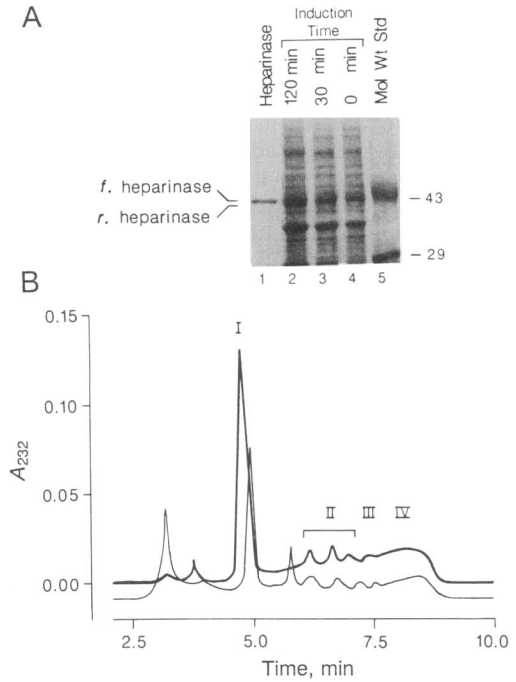

FIG. 3. SDS/PAGE of r-heparinase and *F. heparinum* heparinase and anion-exchange HPLC separation of oligosaccharide products from heparin degraded by heparinase. (*A*) A 12% SDS/PAGE gel of *F. heparinum* heparinase and r-heparinase (*E. coli* whole cell extract). Lane 5, molecular weight standards (expressed as 10^{-3}); lanes 4, 3, and 2, 0, 30, and 120 min, respectively, of induction and r-heparinase production in *E. coli*; lane 1, purified *F. heparinum* heparinase. (*B*) Anion-exchange HPLC separation of heparin oligosaccharides. Separation using a POROS Q/M (4.6 mm × 100 mm) column (PerSeptive Biosystems, MA). ——, *F. heparinum* heparinase-digested heparin fragments; ——, recombinant-digested heparin fragments. I represents the disaccharide, II represents the three tetrasaccharides (T1, T2, and T3), III represents the hexasaccharide, and IV represents the partially digested higher saccharides (27).

within the pRS51hep clone was sequenced in both orientations to yield the 1152-bp ORF of the heparinase I gene. All 11 heparinase I tryptic peptides sequenced mapped onto the ORF (28). The primary protein sequence has a typical prokaryotic leader sequence and cleavage site. Heparinase I gene, without the native leader sequence, was inserted into a pET plasmid for expression. Expression of r-heparinase I lacking the *F. heparinum* leader sequence was successful, and the enzymatic activity of the soluble recombinant protein was similar to that of *F. heparinum* heparinase.

Since heparinase is N-terminally blocked, it has so far not been possible to determine the actual start site. Cysteine labeling and peptide mapping studies clearly indicate the presence of only two cysteines in the *F. heparinum* heparinase I, ruling out the possibility that the signal sequence cleavage site is before a third cysteine residue at position 17 in the signal sequence (R.S., D. Leckband, M.B., C.L.C., and R.L., unpublished observations). Additionally, r-heparinase I, lacking residues 1–21, is still as active as the *F. heparinum* heparinase, indicating that these amino acids are not essential for enzyme activity.

Heparinase activity has been observed in a limited number of prokaryotes. *Bacteroides* sp. (rumen isolates, including *Bacteroides melaningenicus, Bacteroides oralis*, and *Bacteroides ovatus*) show some heparin-degrading activity. Nakamura *et al.* (29) reported recently that a 63-kDa heparinase produced by *Bacteroides heparanolyticus* is catalytically similar to *F. heparinum* heparinase I based on the similarity in heparin degradation products. Even though *Bacteroides* and *Flavobacteria* are phylogenetically related, Southern blotting experiments using the *F. heparinum* heparinase I gene could not detect any cross-hybridizing material in the DNA from *B. heparanolyticus*. It is possible that although similarity may exist at the protein level for the two heparinases, divergence at the DNA level may be too great to detect even in low stringency Southern blots (R.S., unpublished observations).

Heparinase I is an interesting enzyme because of its diverse applications. On the one hand, immobilized heparinase filters, connected to extracorporeal devices, can degrade heparin and neutralize its anticoagulant properties (16). Practical applications of this filter have led to an increased demand for heparinase I, and the production of recombinant heparinase is expected to meet this demand. On the other hand, heparinase I may become a valuable tool in elucidating structure–function relationships of heparin-like complex carbohydrates (28).

We thank Dr. Phil Robbins [Center for Cancer Research, Massachusetts Institute of Technology (MIT)] and Dr. Matt Nugent (Boston University) for reviewing the manuscript. We acknowledge Dr. Joseph Zimmerman for his help and comments during the early stages of this work, Richard Cook (Biopolymers Laboratory, MIT) for peptide sequencing, and Dr. Noubar Afeyan (PerSeptive Biosystems) for the BioCAD station. We thank Ranta Bobba for help in manuscript preparation. This work was supported by funds from the National Institutes of Health (GM 25810 and GM 76454) and the National Science Foundation through the Biotechnology Process Engineering Center, MIT.

1. Kjellen, L. & Lindahl, U. (1991) *Annu. Rev. Biochem.* **60**, 443–475.
2. Jackson, R. L., Busch, S. J. & Cardin, A. D. (1991) *Physiol. Rev.* **71**, 481–539.
3. Klagsbrun, M. & Baird, A. (1991) *Cell* **67**, 229–231.
4. Ruoslahti, E. & Yamaguchi, Y. (1991) *Cell* **64**, 867–869.
5. Comper, W. D. (1981) *Heparin (And Related Polysaccharides)* (Gordon & Breach, New York).
6. Linhardt, R. J., Grant, A., Cooney, C. L. & Langer, R. (1982) *J. Biol. Chem.* **257**, 7310–7313.
7. Vlodavsky, I., Bar-Shavit, R., Ishai-Michaeli, R., Bashkin, P. & Fuks, Z. (1991) *Trends Biochem. Sci.* **16**, 268–271.
8. Linhardt, R. J., Turnbull, J. E., Wang, H.-M., Loganathan, D. & Gallagher, J. T. (1990) *Biochemistry* **29**, 2611–2615.
9. Linhardt, R. J., Galliher, P. M. & Cooney, C. L. (1986) *Appl. Biochem. Biotechnol.* **12**, 135–176.
10. Linker, A. & Hoving, P. (1972) *Methods Enzymol.* **28**, 902–934.
11. Lohse, D. & Linhardt, R. (1992) *J. Biol. Chem.* **267**, 24347–24355.
12. Hoving, P. & Linker, A. (1970) *J. Biol. Chem.* **245**, 6170–6175.
13. Rice, K. G. & Linhardt, R. J. (1989) *Carbohydr. Res.* **190**, 219–233.
14. Linhardt, R. J., Cooney, C. L., Larsen, A., Zannetos, C. A., Tapper, D. & Langer, R. (1984) *Appl. Biochem. Biotechnol.* **9**, 41–55.
15. Yang, V. C., Linhardt, R. J., Bernstein, H., Cooney, C. L. & Langer, R. (1985) *J. Biol. Chem.* **260**, 1849–1857.
16. Langer, R., Linhardt, R. J., Klien, M., Galliher, P. M., Cooney, C. L. & Flannagan, M. M. (1982) in *Biomaterials: Inter-facial Phenomenon and Applications*, eds. Cooper, S., Hoffman, A., Pepas, N. & Ratner, B. (Am. Chem. Soc., Washington, DC), pp. 493–509.
17. Dietrich, C. P., Silva, M. E. & Michelacci, Y. M. (1973) *J. Biol. Chem.* **248**, 6408–6415.
18. Bernstein, H., Yang, V. C., Cooney, C. L. & Langer, R. (1988) *Methods Enzymol.* **137**, 515–529.
19. Cotter, R. (1992) *Anal. Chem.* **64**, 1027–1035.
20. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Plainview, NY).
21. Studier, F. W. (1991) *J. Mol. Biol.* **219**, 37–44.
22. Laemmli, U. K. (1988) *Nature (London)* **227**, 680–685.
23. Moremen, K. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5276–5280.
24. Cardin, A. D. & Weintraub, H. J. R. (1989) *Arteriosclerosis* **9**, 21–32.
25. Von Heijne, G. (1988) *Biochim. Biophys. Acta* **947**, 307–333.
26. Lemp, D., Haselbeck, A. & Klebl, F. (1990) *J. Biol. Chem.* **265**, 15606–15610.
27. Venkataraman, G. (1992) Ph.D. Thesis (Massachusetts Institute of Technology, Cambridge, MA).
28. Sasisekharan, R. (1991) Ph.D. Thesis (Harvard Univ., Cambridge, MA).
29. Nakamura, T., Shibata, Y. & Fujimura, S. (1988) *J. Clin. Microbiol.* **26**, 1070–1071.