

Protein structure and gene cloning of *Syncephalastrum racemosum* nuclease

Heng-Chien HO* and Ta-Hsiu LIAO†¹

*Department of Biochemistry, China Medical College, 91 Hsueh-Shih Road, Taichung, Taiwan, and †Institute of Biochemistry, College of Medicine, National Taiwan University, No. 1, Section 1, Jen-Ai Road, Taipei, Taiwan

The complete amino acid sequence of the fungus *Syncephalastrum racemosum* (Sr-) nuclease has been delineated on the basis of protein sequencing of the intact protein and its protease-digested peptides. The resulting 250-residue sequence shows a carbohydrate side chain attached at Asn¹³⁴ and two half-cystine residues (Cys²⁴² and Cys²⁴⁷) cross-linked to form a small disulphide loop. On the basis of the sequence of Sr-nuclease, a computer search in the sequence database yielded 60% and 48% positional identities with the sequences of *Cunninghamella echinulata* nuclease C1 and yeast mitochondria nuclease respectively, and very little similarity to those of several known mammalian DNases I. Sequence alignment of the three similar nucleases reveals that the single small disulphide loop is unchanged but the carbohydrate attachment in Sr-nuclease is absent from the other two nucleases. Alignment also shows a highly conserved region harbouring Sr-nuclease His⁸⁵, which is assigned as one of the essential residues in the active site. The cDNA encoding Sr-nuclease was amplified by using reverse trans-

criptase-mediated PCR with degenerate primers based on its amino acid sequence. Subsequently, specific primers were synthesized for use in the 3' and 5' rapid amplification of cDNA ends (RACE). Direct sequencing of the RACE products led to the deduction of a 1.1 kb cDNA sequence for Sr-nuclease. The cDNA contains an open reading frame of 320 amino acid residues including a 70-residue putative signal peptide and the 250-residue mature protein. Finally, the recombinant Sr-nuclease was expressed in *Escherichia coli* strain BL21(DE3) in which the recombinant protein, after solubilization with detergent and renaturation, showed both DNase and RNase activities. The assignment of His⁸⁵ to the active site was further supported by evidence that the mutant protein Sr-nuclease (H85A), in which His⁸⁵ was replaced by Ala, was not able to degrade DNA or RNA.

Key words: *Cunninghamella echinulata*, carbohydrate side-chains, cDNA sequences, disulphides, fungi.

INTRODUCTION

Many DNases I from animals [1–8], plants [9] and microorganisms [10–12] have been identified, purified and characterized. Bovine DNase I [13], the most thoroughly studied, is monomeric [14] and so are other mammalian DNases I. However, the fungus *Syncephalastrum racemosum* (Sr-) DNase I [15] exhibits several distinct properties in that the molecule is homodimeric, the specific activity towards duplex DNA is 6-fold that of mammalian DNases I, and the N-terminal 40-residue sequence shows no apparent similarity with those of mammalian DNases I. Hence Sr-DNase I cannot be readily related to mammalian DNases I. Elucidation of the protein and the cDNA sequences for Sr-DNase I was therefore conducted, aiming at a better understanding of the relationships between the Sr- and mammalian DNases I. A sequence homology search in the database indicated, surprisingly, that Sr-DNase I was more closely related to several microbial nucleases than to mammalian DNase I. This finding has prompted us to investigate further the intrinsic RNase activity of Sr-DNase I: the evidence strongly supports the notion that Sr-DNase I is a sugar-non-specific nuclease. Therefore Sr-DNase I, which was named incorrectly, is referred to as Sr-nuclease throughout the rest of this paper.

EXPERIMENTAL

Materials

Baker's yeast RNA, trypsin and chymotrypsin and were purchased from Worthington. Calf thymus DNA, protease V8 (Protease Type XVII), Lys-C endoproteinase, CNBr, Coomassie Brilliant Blue and Cibacron Blue 3GA–Agarose were from Sigma. The 5' rapid amplification of cDNA ends (RACE) system was from Life Technologies. The GeneClean III kit was from BIO 101. The zero-background cloning kit was from InVitrogen. The AmpliCycle sequencing kit for direct sequencing of PCR products was from Perkin-Elmer. Ex Taq DNA polymerase was from TaKaRa. T4 DNA polymerase and T4 DNA ligase were from Epicentre Technologies. Other reagents or chemicals were of the highest grade without further purification.

Purification of Sr-nuclease

The cultivation of *S. racemosum* and the preparation of crude extracts were essentially the same as described previously [15] except that for the last column chromatography Cibacron Blue 3GA–Agarose was used in place of Sephadex G-100. In this dye-affinity chromatography step, the column was equilibrated and

Abbreviations used: DEPC, diethyl pyrocarbonate; rSr-, recombinant *Syncephalastrum racemosum*; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcriptase-mediated PCR; Sr-, *Syncephalastrum racemosum*; Ym-, yeast mitochondrial.

¹ To whom correspondence should be addressed (e-mail thliao@ccms.ntu.edu.tw).

The nucleotide sequence data reported will appear in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession number AF043515.

washed with 20 mM sodium acetate, pH 4.7, before and after sample loading respectively. Sr-nuclease was then eluted with the same buffer containing 1 M $(\text{NH}_4)_2\text{SO}_4$. The purified Sr-nuclease was apparently homogeneous.

DNase I assay

The hyperchromicity assay method [16] was used. One unit of activity was defined as the amount of enzyme causing the increase of one absorbance unit/min at 25 °C in 1 ml of assay buffer, 0.1 M Tris/HCl, pH 7.0, containing 10 mM MnCl_2 , 5 mM CaCl_2 and 50 μg of DNA.

RNase assay

The assay method was essentially similar to that for RNase T1. In brief, each microcentrifuge tube contained 50 μl of 0.2 M Tris/HCl, pH 7.5, 25 μl of 20 mM EDTA, 75 μl of water and 25 μl of enzyme. After the mixture had been prewarmed at 37 °C for 5 min, 62.5 μl of freshly prepared baker's yeast RNA (1.2%, w/v) was added and the incubation continued. After exactly 15 min, 0.1 vol. of 7.5 M ammonium acetate and 3 vol. of ethanol were added and the entire solution was cooled to -20 °C for at least 30 min and centrifuged. The absorbance at 260 nm of the supernatant was measured; one unit of RNase activity was defined as an increase of one absorbance unit.

Protein sequencing

For the preparation of proteolytic digests, 250, 90 or 150 μg of Sr-nuclease were hydrolysed overnight with 5 μg of trypsin, 3 μg of chymotrypsin or 3 μg of Lys-C endoproteinase respectively in 1% (w/v) NH_4HCO_3 at 37 °C. For cleavage at Met residues, 700 μg of Sr-nuclease in 100 μl of 70% (v/v) formic acid was treated with CNBr at room temperature for 24 h. Peptides in each of the resulting digests were separated through a reverse-phase column (Symmetry C_{18} ; Waters) on an HPLC system (Waters); materials from peaks were collected manually and dried in a Speed Vac (Model SC110; Savant). The peptides thus obtained were analysed for amino acid compositions by using phenylisothiocyanate methods [17] and sequenced on a protein sequencer (Model 477A; Applied Biosystems).

Sequence similarity

The entire 250 residues of Sr-nuclease were used as the query sequence to search for sequence similarity by using PC/Gene and BLAST (Basic Local Alignment Search Tool). The scan was done using a k -tuple value of 1. The distance parameter was set to 2 amino acids. The final scores were computed by using the Dayhoff MDM-78 matrix. The scan was performed with all the protein sequences in the database CDPROT29. The results were: number of sequences successfully scanned, 38303; average score in this scan with the current parameters, 200.2417; standard deviation, 52.29214; reference score of the sequence against itself, 7282.

Total RNA isolation

S. racemosum was cultivated in a medium containing 0.8% agarose, 0.1% glucose and 20% (w/v) wheat bran at 30 °C for 5 days. Multiplied mycelia were then carefully removed and homogenized with TRIzol reagent. The homogenate was immediately temperature-treated at 60 °C for at least 30 min to inactivate Sr-RNase. Total RNA was isolated from the homogenate with TRIzol Reagent (Life Technologies) in accordance with the manufacturer's protocol.

Reverse transcriptase-mediated PCR (RT-PCR)

Total RNA (3.92 μg), 0.5 μg of oligo(dT) primer and diethyl pyrocarbonate (DEPC)-treated water were combined to a final volume of 12.5 μl and cDNA templates were synthesized with first-strand cDNA synthesis kit (Novagen) in accordance with manufacturer's instructions. To amplify cDNA of Sr-nuclease, the components in 50 μl of solution were as follows: 1.5 mM MgCl_2 , 0.2 mM dNTP mix, 300 pmol of 5'-GCDATYTTTAAARAARTG-3' [(anti-sense, degenerate, Sr-nuclease(179–184))], 300 pmol of 5'-CAYTGGGTNGGNGARCA-3' [(sense, degenerate, Sr-nuclease(35–40))] or 5'-GAYATHCCNGARATGTT-3' [(sense, degenerate, Sr-nuclease(64–69))], cDNA templates (equivalent to approx. 200 ng of total RNA) and 2.5 units of Ex Taq. The PCR consisted of one initiation cycle (5 min at 95 °C), 35 amplification cycles (1 min at 95 °C, 1 min at 50 °C and 1 min at 72 °C) and one termination cycle (7 min at 72 °C). Before DNA sequencing, the RT-PCR products were purified with a GeneClean III kit. DNA sequencing was performed on a DNA sequencer (Model 377, Perkin-Elmer).

3' RACE

Total RNA (0.78 μg), 10 pmol of the adapter primer (5'-GGC-CACGCGTGCAGTAGTACT₁₇-3') and DEPC-treated water were combined in a final volume of 12 μl . Before proceeding with reverse transcription, as suggested in the manufacturer's instructions (Life Technologies), the mixture was heated at 70 °C for 10 min and chilled on ice for at least 1 min. When the primary target cDNA was being amplified, the components in 50 μl of solution were as follows: 1.5 mM MgCl_2 , 0.2 mM dNTP mix, 10 pmol of universal amplification primer (5'-CUACUACUACUAGGCCACGCGTGCAGTAGTAC-3'), 10 pmol of 5'-CAA-GAAGCCATGGACGAG-3' [sense, specific, Sr-nuclease(96–101)], 2 μl of cDNA templates and 2.5 units of Ex Taq. The PCR was performed in the following manner: initiation at 95 °C for 5 min, 35 cycles of amplification (1 min at 95 °C, 1 min at 50 °C and 2 min at 72 °C) and final extension at 72 °C for 7 min. For nested amplification, the components in 50 μl of solution were as follows: 1.5 mM MgCl_2 , 0.2 mM dNTP mix, 10 pmol of the abridged universal amplification primer (5'-GGCCACGCGTGCAGTAGTAC-3'), 10 pmol of 5'-ACCTTTTGTCTCTCCA-AT-3' [sense, specific, Sr-nuclease(102–107)], 2 μl of the above primary PCR products and 2.5 units of Ex Taq. The PCR conditions for amplifying nested target cDNA were the same as those for the primary PCR. The nested product of 3' RACE was purified with GeneClean III kit and then subjected to DNA sequencing.

5' RACE

Total RNA (0.78 μg), 2.5 pmol of 5'-CTTTCGACCCGTGCTCGC-3' [anti-sense, specific, Sr-nuclease(151–156)] and DEPC-treated water were combined in a final volume of 15.5 μl . The mixture was incubated at 70 °C for 10 min and chilled on ice for 1 min. The procedure in the subsequent steps were as given in the manufacturer's instructions, including the synthesis of first-strand cDNA, the spin cartridge purification of cDNA and the dCTP-tailing of cDNA by terminal transferase. The components of the primary PCR reaction were, in 50 μl of solution: 1.5 mM MgCl_2 , 0.2 mM dNTP mix, 20 pmol of 5'-AGGACCGGTGT-AAACATA-3' [anti-sense, specific, Sr-nuclease(139–144)], 20 pmol of abridged anchor primer (5'-GGCCACGCGTGCAGTAGTACGGGIIIGGGIIGGGIIG-3'), 5 μl of dC-tailed cDNA and 2.5 units of Ex Taq. The amplification conditions were the same as those for the 3' RACE primary PCR reaction except for

a final extension at 72 °C for 1 min. For re-amplification of the target cDNA, the components were, in 50 µl of solution: 1.5 mM MgCl₂, 0.2 mM dNTP mix, 20 pmol of 5'-GAACATCTCGG-GAATATC-3' [anti-sense, specific, Sr-nuclease(64–69)], 20 pmol of abridged universal amplification primer, 5 µl of the primary PCR product and 2.5 units of Ex Taq. PCR was initiated at 95 °C for 2 min; there followed 35 cycles of amplification (1 min at 95 °C, 1 min at 46 °C and 1 min at 72 °C) and extension for 7 min at 72 °C. After filling-in at both ends by T4 DNA polymerase and purification with GeneClean III kit, the secondary PCR product of 5' RACE was subcloned into the vector pZerO-2.1 at the *EcoRV* site and transformed into TOP10F'-competent cells. The recombinant plasmid was cycle-sequenced with M(-13) forward primer (5'-GTAAAACGACGGCCAG-3') and reverse primer (5'-CAGGAAACAGCTATGAC-3').

Construction and expression of recombinant *S. racemosum* (rSr-) nuclease

The full-length cDNA encoding the mature Sr-nuclease was amplified by using PCR. The components in 50 µl of solution were as follows: 1.5 mM MgCl₂, 0.2 mM dNTP mix, 30 pmol of primer EB-MASSDI (5'-CACTTGAATTCGGATCCATGGC-ATCCTCCGATATCC-3') containing a *Bam*HI site, 30 pmol of primer VR-EPS (5'-CCCGGGCTGCAGAATTCAATATCATCGTAC-3') containing an *Eco*RI site, cDNA templates [approx. 200 ng of total RNA and reverse-transcribed from oligo(dT) primer] and 2.5 units of Ex Taq. The PCR reaction was initiated at 95 °C for 5 min; there followed 35 cycles of amplification (1 min at 95 °C, 1 min at 50 °C and 1 min at 72 °C) and extension at 72 °C for 7 min. After being digested with *Eco*RI and *Bam*HI, the PCR product was subcloned into the plasmid pET-29(a)⁺. The resulting recombinant plasmid, termed pFU.DNase2, was maintained in *Escherichia coli* strain DH5α. The following protocols for gene expression were equivalent to those provided by the manufacturer (Novagen). The expression host, *E. coli* strain BL21(DE3) harbouring pET-29(a)⁺ or the recombinant plasmid pFU.DNase2, was cultured at 37 °C in 40 ml of Luria-Bertani medium until D₆₀₀ reached 0.6 or higher. The final concentration of isopropyl β-D-thiogalactoside for induction was 1 mM. At each duration of induction, an aliquot of bacterial cultures was harvested and lysed by sonication in 1% (w/v) SDS. After centrifugation, the clear supernatant was analysed by SDS/PAGE.

PCR mutagenesis

Residue His⁸⁵ of Sr-nuclease was replaced by Ala with the use of PCR mutagenesis. The region encoding residues 1–88 of mature Sr-nuclease was first PCR-amplified by using plasmid pFU.DNase2 as template. Two primers were EB-MASSDI (see above) and H85A-RP (5'-AGGCGCTTGC GCGCCACGGTC-ATAGCCACT-3') containing a *Bss*HII site. Another region encoding residues 83–250 of Sr-nuclease was also PCR-amplified under the same conditions except that primers VR-EPS (see above) and H85A-FP (5'-CGTGGCGCGCAAGCGCCTGCTGCCGATGAT-3') containing a *Bss*HII site were used. Subsequently, both PCR products were mixed for the secondary PCR reaction with primers EB-MASSDI and VR-EPS. Thus the product of secondary PCR reaction contained an alanine residue in place of His at position 85. After being digested with *Bam*HI and *Eco*RI, the PCR product was subcloned into the corresponding sites of pFU.DNase2 to exchange the insert. The resulting recombinant plasmid, termed pFU.DNase4, was maintained in *E. coli* strain NovaBlue. Expression of the mutant

plasmid was exactly the same as that of the wild type, pFU.DNase2.

Zymogram methods

The zymogram methods for DNase I and RNase assays were essentially those described by Rosenthal and Lacks [18], with minor modifications. Calf thymus DNA or baker's yeast RNA were included, when preparing the separating gels, to a final concentration of 15 or 80 µg/ml respectively. After being mixed with SDS-containing sample buffer, the samples were heated in boiling water for 5 min. After electrophoresis, the gel was soaked at room temperature in 75 ml of 10 mM Tris/HCl, pH 7.0, containing 1 mM MgCl₂ for 30 min with gentle shaking. This process was repeated twice. Finally, the gel was incubated with the same buffer at 37 °C for 30 min and stained with ethidium bromide. The DNase or RNase activities showed as dark bands under UV irradiation.

RESULTS

Chemical structure of Sr-nuclease

Figure 1 shows the covalent structure of Sr-nuclease. The polypeptide chain consists of 250 residues in which the two half-cystine residues (Cys²⁴² and Cys²⁴⁷) are cross-linked to form a small disulphide loop. This therefore confirms the previous finding [15] that the homodimeric Sr-nuclease molecule is the result of non-covalent interactions and not through the disulphide linkages between the two monomers.

Sr-nuclease is a glycoprotein [15]; the amino acid sequence (Figure 1) shows a potential N-glycosylation site at Asn¹³⁴. When peptide Cb7 and peptide T9 were sequenced, in the sixth and the fifth cycles respectively, neither phenylthiohydantoin-Asn nor any other identifiable phenylthiohydantoin-amino acid appeared, indicating that a carbohydrate side chain was attached to Asn¹³⁴.

Sequence similarity

Based on the sequence of Sr-nuclease, the computer search in the sequence database yielded the highest similarity scores with *Cunninghamella echinulata* nuclease C1 [19] and yeast mitochondrial (Ym-) nuclease [20]. Figure 2 shows a sequence alignment for the three nucleases. Sr-nuclease shared 60% (149/250 residues conserved) and 48% (120/250 residues conserved) identities with nuclease C1 and Ym-nuclease respectively. If one separates the polypeptide chain into two parts, the conserved residues are distributed mainly in the N-terminal 150 residues of Sr-nuclease; here the identity increased to 55–63% (82–95/150 residues conserved). In contrast, only 38–54% identity was observed in the C-terminal 100 residues of Sr-nuclease (38–54/100 residues conserved). Two regions (Ser¹⁴⁹–Pro¹⁷² and Ile¹⁸³–Glu¹⁹² of Sr-nuclease) show the most divergence. It is noteworthy that there are 54 and 25 residues protruding at the N- and C-termini respectively of the polypeptide chain deduced from the cDNA sequence of Ym-nuclease.

The alignment also shows several important features. The single small disulphide loop in all three nucleases is unchanged but the carbohydrate attachment site in Sr-nuclease is absent from the other two nucleases. The region harbouring residue His⁸⁵ of Sr-nuclease is the most conserved, showing consecutive sequences (SGXDRGHXAPAXD). A similar sequence (Figure 2, line 4) can also be found in *Serratia marcescens* nuclease [21], another microbial nuclease, for which the three-dimensional structure has been determined [22] and for which the His residue (His⁸⁹) in the sequence was indicated to be at the active site [23]. Thus, by similarity, His⁸⁵ of Sr-nuclease is likely to be at the



Figure 1 Complete amino acid sequence of Sr-nuclease

Peptides obtained by chemical and enzymic cleavages of the intact Sr-nuclease are mapped under the sequence. The prefixes T, L, C, Cb and V represent peptides from the digests of trypsin, Lys-C endoproteinase, chymotrypsin, CNBr and protease V8 respectively. Daughter peptides are designated with letters following the names of the parental peptides. The numbers at the right of each line indicate the residue number of the last amino acid residue of that line. Solid lines indicate sequences completed by Edman degradation; broken lines, indicate sequences completed by composition only; CHO and -S-S- indicate the carbohydrate side chain and the disulphide bridge respectively. Both peptide L21 and peptide C21 contained an intra-chain disulphide linkage. Peptide T12-13, containing two peptides linked by a disulphide bridge, was eluted in a single fraction on HPLC and the results from the protein sequencer showed two peptide sequences for this fraction. This protein sequence has been deposited in the SWISS-PROT data bank under the accession number P81204.

Ym-Nuclease	MCSRILLSGVLGAGTGLTYLLLNKHSPTQIIETYPPTQKFNISNIQSH	50
Sr-Nuclease		
Nuclease C1		
Sm-Nuclease		D 1
Ym-Nuclease	SFNVDPGSGFFKYGFPGPIHDLQNRREEFISCYNROTQNPYVWLEHITPESEL	100
Sr-Nuclease	ASSDILKLGNGPVPVSDLLERSGYILSYNRRDRLAHVWGEHLTSASL	46
Nuclease C1	SPNSESILQFQDEGTARDFLERESYVSYNRRDRVASWTGEHLTADSL	48
Sm-Nuclease	TLESIDNCVGCPTGGSSNVSVIRHAYTLNNSSTTKFANVWYVHITKDTF	51
Ym-Nuclease	AARNA-DRKNSEFFKEDVEIPEKFRGKLRDY----FRSGYDRGHQAPAAAD	145
Sr-Nuclease	QAGQGVDRDKSNFOEDTDIPEMFRHLKDY----VSSGYDRGHQAPAAADD	92
Nuclease C1	KTGDGVDRDHSKFKEDDVEESLERSLADY----SGSGFDRGHMAPAGDA	94
Sm-Nuclease	ASGK--TRN---WKTDEALNPADTLAPADYTGANAALKVDRGHQAPLASL	96
Ym-Nuclease	KFSQOAMDDTEFLSNMCPVGVGFNRDYYWAHLEYFCRGLTKKVK--SVRI	193
Sr-Nuclease	LSSQFAMDETEFLSNMCPVGVGFNRDYYWAHLEYFCRGLTKKVK--SVRI	140
Nuclease C1	VATQFAMDDTEFLSNMCPVGVGFNRDYYWAHLEYFCRGLTKKVK--SVRI	142
Sm-Nuclease	AGVSD-WESLNLYLSNITPQKSD-LNQGAWARLEDCERKIDRADISSVYT	144
Ym-Nuclease	VTGPLY--LPKIDPINDK--FRVNYEYIGN-PPSIAVPTHEFEKLVIAEA	237
Sr-Nuclease	YTGPLE--LPSAASTGRKNPAYSIEYFPLGATTPNVPVTHPEK--IALT	183
Nuclease C1	FTGPLE--LPTKGSDEK--YIVTYNVEQG--NVAVPTHFYK--VILV	178
Sm-Nuclease	VTGPLYERDMGKLPQTKR-----AHTIPSAIYKVIIFINN	178
Ym-Nuclease	PTANPAREDIRVAAFVLPNRPISNETKLTDFEVPIDALERSTGEELLOKV	295
Sr-Nuclease	TTA---SSFYALGAEVYLPNRIIDSTPLTNEKVELEAIEKAAGLVFFDKL	241
Nuclease C1	PQG---DNKYAYGAEVYLPNRIIDSTPLTNEKVELEAIEKAAGLVFFDKL	236
Sm-Nuclease	SFA-----VNHYAALF-DQNTPKGADFQERVTVDIEIKRPTGLIIWAGL	222
Ym-Nuclease	EPKSKKALCKEIVNCQIVVDRFSNAAIKQSKDVKLLPFPKKN	329
Sr-Nuclease	DRSKFADLCSKTCQVR	250
Nuclease C1	DVSTLGLDECATGAVSSSGGGDA	252
Sm-Nuclease	DDDVQASLKKKPGVLEFELMGCKN	245

Figure 2 Amino acid sequence alignment

Numbers at the right relate to the last residue on each line. The numbering for the sequence of Ym-nuclease is the same as that deduced from the cDNA sequence [20]. Gaps are inserted as dashes to gain maximum similarity. Identical residues are shaded. The colon indicates the active-site His residue. Abbreviation: Sm-, *Serratia marcescens*.

active site as well. Furthermore, the conserved DR dipeptide sequence was shown to be catalytically relevant in *S. marcescens* nuclease [23].

Intrinsic RNase activity

In view of sequence similarity with other nucleases as illustrated in Figure 2, one cannot eliminate the possibility that Sr-nuclease

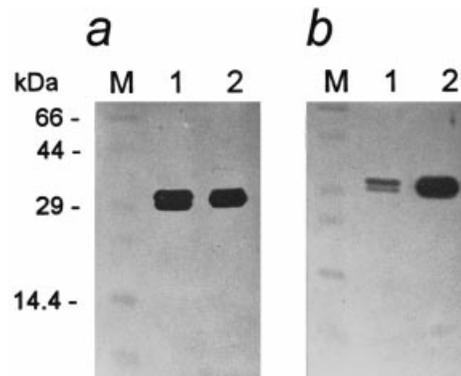


Figure 3 Zymograms for DNase and RNase assays

Enzymes were first precipitated by the addition of an equal volume of 10% (w/v) trichloroacetic acid. The pellets were resuspended in SDS-containing sample buffer to a final concentration of 0.2 µg/µl; the solutions were heated in boiling water for 5 min before being loaded. (a) Zymogram for DNase assay; (b) zymogram for RNase assay. In both panels, lanes 1 and 2 contained purified Sr-nuclease (0.4 µg) and nuclease C1 (0.4 µg) respectively; lanes M contained prestained protein markers (molecular masses indicated in kDa at the left).

has the ability to degrade RNA. Therefore the intrinsic RNase activity for Sr-nuclease was examined with the use of the zymogram method [18], one of the most sensitive methods for RNase assay. The results showed that, in addition to the DNase activity (Figure 3a, lane 1), Sr-nuclease did indeed possess RNase activity (Figure 3b, lane 1). For comparison, the well-established nuclease C1 [19] showed similar results (Figures 3a and 3b, lanes 2). On the basis of the intensities of the active bands

```

                                     ATCCCTTCTTTCTCT 16
TCITGCTTTCCCGCTCTTTCTTTCTTATCTCTCTGTGCTGCTGCTTTCTTTGCG 76
ATGCCATCATCAGGAGAACCGACAGACTACTGCCAGCAGCAGTAAAGTGCACCTTCC 136
M P I I R R T S R A T A S S S K S A T S
AGTATCTGCGACAGGCCATTCTTGTCTTCTACTGGCTTTCTGCTGGTGCACTCTTCA 196
S I L R Q A I L V F T G F L L G A L F M
GTCTTTGTCCTTCGTTGTATAATGCCCGAACCATCGTCTCCATCGGTCATGAAT 256
V F V P S L Y N A A E P S S L H R L M N
CCCATTCGGCGCAACACCCTCAGCGCACGTGATCCTCCGATATCCTTAAATGGGCAAT 316
P I R R N T L S A R A S S D I L K L G N
CCAGGACCGGTGAGCGATCTCTTGAACGATCCGGTATATTCTGTCTTACAATCGACG 376
P G F V S D L L E R S G Y I L S Y N R R
GACCGACTGGCGCACTGGGTGGTGAACATCTCAGGAGCGCTTCGTTGCAAGCCGGCC 436
D R L A H W V G E H L T S A S L Q A G Q
GGCGTGATCGTGATAAATCCAACITTCAGGAGGACACTGATATCCCGAGATGTTCCGT 496
G V D R D K S N F Q E D T D I P E M F R
GCTCATCTCAAGACTAGCTGAGCAGTGGCTATGACCGTGGCCATCAAGCGCTGCTGCC 556
A H L K D Y V S S G Y D R G H Q A P A A
GATGATTTATCCTCTCAAGAAGCCATGGACGAGACCTTTTGTCTTCCAATATGGCCCT 616
D D L S S Q E A M D E T F L L S N M A P
CAAGTCGGCGTGGGTTCAATCGCCACTATTGGGCTACCTCGAGGGTTTCATGCGAGAC 676
Q V G V G F N R H Y W A Y L E G F M R D
TTGACCCAGAAATTCAGTACGCTGTATGTTTACCCGGTCTCTTTTCCGCTCAGCC 736
L T Q N F T D V Y V Y T G P L F L P S A
GCGAGCAGGGTTCGAAAGAACCAGGCTTATAGCATCGAATATCCTTCCCTCGGTGCTACC 796
A S T G R K N P A Y S I E Y P F L G A T
ACACCCAAATGTTCCGGTGGCGACCCACTTCTTCAAGATCGCCCTTGACAACCAGGCCCTC 856
T P N V P V P T H F F K I A L T T T A S
TCGGAATATGCCCTAGGTGCCCTTTGTGTTACCGAACCGGCTATCGACAGTAGCACC 916
S E Y A L G A F V L P N Q A I D S S T P
CTCACCAATTTAAGGTGGAGCTCGAAGCCATTGAAAAGCAGCTGGTCTCGTCTTTTT 976
L T N F K V E L E A I E K A A G L V F F
GACAAGCTTGACCGCTCAAGTTTGGCGACTTGTGCTCTAAAACCACCTTGTCAAGTAG 1036
D K L D R S K F A D L C S K T T C Q V R
TGATATTGAATAAAGTACATAGGATGTGTGACGTTTAAATAGTCCGACTGATATGATATA 1096
*
CAAAAAATAAATA 1110

```

Figure 4 Sequence for a cDNA encoding Sr-nuclease

Numbers at the right relate to the last nucleotide on each line. The translated protein sequence is shown under the nucleotide sequence. The asterisk indicates the stop codon.

for the same quantities of proteins loaded, Sr-nuclease and nuclease C1 had approximately the same level of DNase activities but Sr-nuclease was less active than nuclease C1 in RNase assay (Figure 3b, lanes 1 and 2). The intrinsic RNase activity of Sr-nuclease was also examined by the standard RNase assay method normally used for measuring fungal RNase T1 activities. The results showed that the purified Sr-nuclease had a specific activity of 21 units/mg, which was also lower than that of the purified nuclease C1 (36 units/mg), consistent with the results obtained by the zymogram assay (Figure 3b).

cDNA sequence

On the basis of the amino acid sequence of Sr-nuclease (Figure 1), three degenerate primers were designed to amplify the cDNA of Sr-nuclease. Two RT-PCR products of 450 and 360 bp were obtained by using two combinations of two upstream sense primers (corresponding to residues 35–40 and 64–69) and one downstream anti-sense primer (corresponding to residues 179–184). Direct sequencing positively identified both RT-PCR products as the polynucleotides amplified from the cDNA of Sr-nuclease. Two specific upstream and three specific downstream primers were then synthesized for use in the 3' and 5' RACEs.

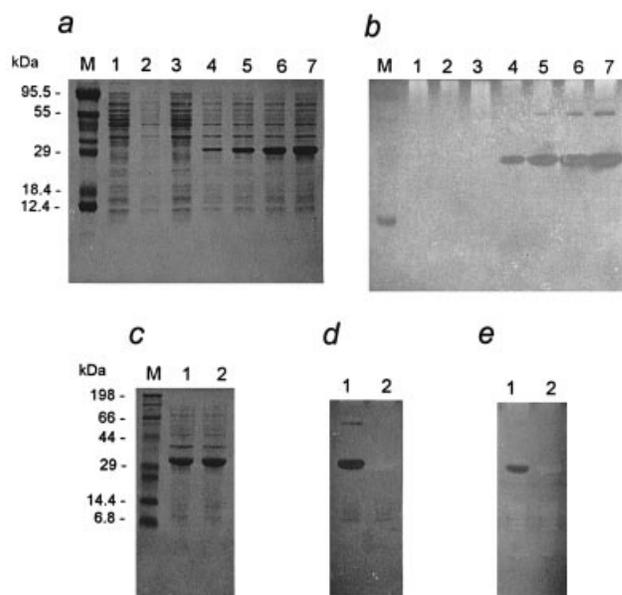


Figure 5 Expression products analysed by SDS/PAGE

(a, c) Gels stained with Coomassie Brilliant Blue; (b, d) zymograms for DNase assay; (e) zymogram for RNase assay. *E. coli* cells harbouring plasmids were harvested and lysed by adding 0.1 vol. of 1% (w/v) SDS. In (a) and (b), lanes 1 and 2 contained the lysates of pET-29(a)⁺ vector-transfected control cells, 0 and 3 h respectively after induction; lanes 3–7 contained the lysates of the pFU.DNase2-transfected cells, 0, 0.5, 1, 2 and 3 h respectively after induction. In (c), (d) and (e), lanes 1 and 2 contained the lysates of the pFU.DNase2- and pFU.DNase4-transfected cells respectively, 3 h after induction. In all panels lanes M contained pre-stained protein markers (molecular masses indicated in kDa at the left).

The resulting PCR products from both RACEs were approx. 650 and 500 bp in length. From the sequences of these 3' and 5' RACE products the cDNA nucleotide sequence for Sr-nuclease was derived (Figure 4). The cDNA was a full-length DNA of approx. 1.1 kb, containing an open reading frame encoding 320 amino acid residues. The 5' and 3' untranslated regions were 74 and 71 bp respectively. The open reading frame consisted of 70 residues of a putative signal peptide and 250 residues of the mature Sr-nuclease. The entire protein sequence of Sr-nuclease shown in Figure 1 matched perfectly with that deduced from the cDNA nucleotide sequence in Figure 4. Three in-frame Met residues were found in the signal peptide. The N-terminal Ala of the mature protein began with nt 285–287 (GCA) and the C-terminal Arg ended with nt 1034–1036 (CGA), immediately followed by the termination codon (TGA).

Expression of rSr-nuclease

The cDNA encoding the mature Sr-nuclease of 250 residues as shown in Figure 4 was amplified by PCR. The recombinant plasmid pFU.DNase2 was constructed in which the PCR-amplified DNA fragment plus the vector-encoded 33 residues at the N-terminus (total 283 residues of encoded DNA) was inserted into the expression vector pET-29(a)⁺ at the *Bam*HI and *Eco*RI sites. Figure 5(a) shows the results of expression of pFU.DNase2 in *E. coli* strain BL21(DE3) at 37 °C. The rSr-nuclease protein was not found in the growth medium but was detectable within 30 min of induction (Figure 5a, lane 4) in the cell lysate solubilized with detergent. The amount of rSr-nuclease 3 h after induction was estimated to be more than 10 mg on the basis of the SDS/PAGE results (Figure 5a, lane 7) for 40 ml of culture. In

contrast, there was little or no change in the protein in the vector-only control group after induction (Figure 5a, lanes 1 and 2). The expression efficiency was severely affected when cells were cultured at lower temperatures (30 °C and below). The DNase activity of rSr-nuclease determined by the zymogram method is shown in Figure 5(b). The removal of detergent and renaturation restored the activity of the recombinant protein to degrade DNA (Figure 5b, lanes 3–7). There was little or no detectable DNase activity in the control cell lysates (Figure 5b, lanes 1 and 2). Because the active rSr-nuclease is not a glycoprotein, the extra carbohydrate side chain in the native Sr-nuclease, isolated from the fungal extracts, must not be a required moiety for activity. By the same reasoning, the extra N-terminal 33 residues in rSr-nuclease also exert no influence on enzyme activity.

Active site of Sr-nuclease

The sequence alignment shown in Figure 2 implies that His⁸⁵ of Sr-nuclease is at the active site. This was confirmed by PCR mutagenesis by altering His⁸⁵ to Ala. The mutant protein, rSr-nuclease(H85A), was produced in *E. coli* strain BL21(DE3)pLysS with the mutant plasmid pFU.DNase4. The expression efficiencies were about the same for both the wild type and the mutant (Figure 5c, lanes 1 and 2). However, the DNase and RNase activities were completely abolished in the mutant (Figures 5d and 5e, lanes 2) in contrast with the wild type (Figures 5d and 5e, lanes 1). The results indicate that His⁸⁵ of Sr-nuclease is the active-site residue and both the DNase and RNase activities are conferred by the same His⁸⁵.

DISCUSSION

On the basis of sequence similarity, Ym-nuclease, Sr-nuclease and nuclease C1 should be categorized into a new family that should also include at least two other nucleases from prokaryotic members, namely *Anabaena* sp. PCC 7120 [24] and *S. marcescens* [21]. In contrast, other prokaryotes, such as *E. coli* [25], *Aeromonas hydrophila* [26], *Vibrio cholerae* [27] and *Erwinia chrysanthemi* [28], produce sugar-specific DNases of entirely different amino acid sequences. Therefore on the basis of sequence similarity it is possible to redefine a previously known DNase in *S. racemosum* as a sugar-non-specific nuclease. In this respect, the sequences of the nucleases secreted by *Aspergillus nidulans* [11] could be investigated for a better understanding of their biological functions.

On several occasions with the RNase assay method normally used for RNase T1, the substrate, RNA, was readily hydrolysed by the crude Sr-nuclease extracts but not by the purified Sr-nuclease (results not shown). It was therefore concluded that Sr-nuclease had no RNase activity and RNA was degraded by the specific Sr-RNase in the crude extracts. In fact, Sr-RNase has been purified and, on the basis of its molecular mass and partial amino acid sequence, it falls into the category of the RNase T2 family (C.-C. Ho, H.-C. Ho and T.-H. Liao, unpublished work). However, the possibility the intrinsic RNase activity in Sr-nuclease still cannot be excluded on the basis of sequence similarity with other microbial nucleases (Figure 2). This speculation was finally confirmed by the sensitive zymogram assay method (Figure 3b) and further demonstrated by using the assay method for RNase T1 with a larger amount of enzyme.

Ym-nuclease is located in the inner membrane of mitochondria [29]. The cloned Ym-nuclease gene showed an encoded protein of 329 residues with a calculated molecular mass of 37 kDa, very close to that of the Ym-nuclease (38 kDa). Therefore the prosequence of the Ym-nuclease precursor was very short or uncleavable. In contrast, the putative signal peptide of Sr-

nuclease was extremely long and similar to prepro sequences of some aspartic proteinase precursors [30,31]. Because the sequence from Met⁻⁷⁰ and forward has the characteristics of signal peptides [32], it possibly represents the translational start.

The N-terminal protruding region of Ym-nuclease (Figure 2) was searched for sequence similarity with all of the RNase family by using the BLAST program. The results showed that such a region and the first halves of *Aspergillus oryzae* RNase T1 [33] and of *Neurospora crassa* RNase N1 [34] shared a smaller extent of identity, implying that this region is possibly not related to RNase activity. There is other evidence to support this suggestion. First, Sr-nuclease and nuclease C1 have approximately the same molecular masses and similar sequences, as illustrated in Figure 2. Secondly, *S. marcescens* nuclease, containing 245 residues, also has a similar size, as does Sr-nuclease, and lacks the N-terminal protruding region of Ym-nuclease. Therefore the regions showing the highest degree of identity between Sr-nuclease and Ym-nuclease should be those involved in nuclease activity. In contrast, the regions Ser¹⁴⁹-Pro¹⁷² and Ile¹⁸³-Glu¹⁹² of Sr-nuclease (Figure 2) are the most divergent among nucleases of the same family and are irrelevant to nuclease activity.

The production of Sr-nuclease was inducible and the enzyme was secreted into the extracellular medium as opposed to Ym-nuclease, which is targeted to mitochondria. The distinct localizations of these enzymes might contribute to the differences in signal peptides [20]. Zassenhaus and Denniger [35] reported that Ym-nuclease was involved in DNA recombination and gene conversion within mitochondria. The physiological function of the secreted Sr-nuclease is not known but could be very different from that of Ym-nuclease. The role of the secreted Sr-nuclease might not simply be the degradation of DNA to provide nutrients for fungal growth.

This work was supported in part by grants from the National Science Council of the Republic of China (NSC86-2341-B-002-039 to T. L. and NSC87-2314-B-040-019 to H. H.).

REFERENCES

- Price, P. D., Liu, T.-Y., Stein, W. H. and Moore, S. (1969) *J. Biol. Chem.* **244**, 917–923
- Ball, W. D. (1974) *Biochim. Biophys. Acta* **341**, 305–317
- Murai, K., Yamanaka, M., Akagi, K., Anai, M., Mukai, T. and Omae, T. (1978) *Biochim. Biophys. Acta* **517**, 186–194
- Wadano, A., Hobus, P. A. and Liao, T.-H. (1979) *Biochemistry* **18**, 4124–4130
- Paudel, H. K. and Liao, T.-H. (1986) *J. Biol. Chem.* **261**, 16006–16011
- Chou, M.-Y. and Liao, T.-H. (1990) *Biochim. Biophys. Acta* **1036**, 95–100
- Yasuda, T., Takeshita, H., Nakajima, T., Hosomi, O., Nakashima, Y. and Kishi, K. (1997) *Biochem. J.* **325**, 465–473
- Hsiao, Y.-M., Ho, H.-C., Wang, W.-Y. and Liao, T.-H. (1997) *Eur. J. Biochem.* **249**, 786–791
- Liao, T.-H. (1977) *Phytochemistry* **16**, 1469–1471
- Lehman, I. R., Roussos, G. G. and Pratt, E. A. (1962) *J. Biol. Chem.* **237**, 819–828
- Campbell, A. M. and Winder, F. G. (1983) *Biochim. Biophys. Acta* **746**, 125–132
- Li, C.-J., Hwa, K.-Y. and Englund, P. T. (1995) *Nucleic Acids Res.* **23**, 4426–4433
- Moore, S. (1981) in *The Enzymes*, 3rd edn., vol. 14A2, (Boyer, P. D., ed.), pp. 281–296, Academic Press, New York
- Lindberg, U. (1967) *Biochemistry* **6**, 335–342
- Chen, L.-Y., Ho, H.-C., Tsai, Y.-C. and Liao, T.-H. (1993) *Arch. Biochem. Biophys.* **303**, 51–56
- Liao, T.-H. (1974) *J. Biol. Chem.* **249**, 2354–2356
- Heinrikson, R. L. and Meredith, S. C. (1984) *Anal. Biochem.* **136**, 65–74
- Rosenthal, A. L. and Lacks, S. A. (1977) *Anal. Biochem.* **80**, 76–90
- Ho, H.-C., Shiau, P.-F., Liu, F.-C., Chung, J.-G. and Chen, L.-Y. (1998) *Eur. J. Biochem.* **256**, 112–118
- Vincent, R. D., Hofmann, T. J. and Zassenhaus, H. P. (1988) *Nucleic Acids Res.* **16**, 3297–3312
- Ball, T. K., Saurugger, P. N. and Benedik, M. J. (1987) *Gene* **57**, 183–192
- Lunin, V. Y., Levnikov, V. M., Shlyapnikov, S. V., Blagova, E. V., Lunin, V. V., Wilson, K. S. and Mikhailov, A. M. (1997) *FEBS Lett.* **412**, 217–222

- 23 Friedhoff, P., Gimadutdinov, O. and Pingoud, A. (1994) *Nucleic Acids Res.* **22**, 3280–3287
- 24 Muro-Pastor, A. M., Flores, E., Herrero, A. and Wolk, C. P. (1992) *Mol. Microbiol.* **6**, 3021–3030
- 25 Jekel, M. and Wackernagel, W. (1995) *Gene* **154**, 55–59
- 26 Chang, C.-C., Chang, S.-Y., Chen, S.-L. and Chuang, S.-M. (1992) *Gene* **122**, 175–180
- 27 Focareta, T. and Manning, P. A. (1987) *Gene* **53**, 31–40
- 28 Moulard, M., Condemine, G., Nasser, W. and Robert-Baudouy, J. (1995) *Biochim. Biophys. Acta* **1262**, 133–138
- 29 Dake, E., Hofmann, T. J., McIntire, S., Hudson, A. and Zassenhaus, H. P. (1988) *J. Biol. Chem.* **263**, 7691–7702
- 30 Berka, R. M., Carmona, C. L., Hayenga, K. J., Thompson, S. A. and Ward, M. (1993) *Gene* **125**, 195–198
- 31 Runeberg-Roos, P., Tormakangas, K. and Ostman, A. (1991) *Eur. J. Biochem.* **202**, 1021–1027
- 32 von Heijne, G. (1985) *J. Mol. Biol.* **184**, 99–105
- 33 Takahashi, K. (1985) *J. Biochem.* **98**, 815–817
- 34 Takahashi, K. (1988) *J. Biochem.* **104**, 375–382
- 35 Zassenhaus, H. P. and Denniger, G. (1994) *Curr. Genet.* **25**, 142–149

Received 16 November 1998/7 January 1999; accepted 21 January 1999