Cloning and characterization of a novel nuclease from shrimp hepatopancreas, and prediction of its active site

Wen-Yi WANG*, Shwu-Huey LIAW† and Ta-Hsiu LIAO*†

*Institute of Biochemistry, College of Medicine, Taiwan University, 1, Sec. 1, Jen-Ai Road, Taipei, Taiwan, and †Department of Life Science, National Yang-Ming University, 155, Sec. 2, Li-Nong Street, Taipei, Taiwan

INTRODUCTION

Nucleases, in a broad sense, are enzymes capable of hydrolysing nucleic acids, including not only DNase and RNase but also the sugar-non-specific nucleases that have been named nucleases for short. DNase I, cleaving double-stranded DNA to yield 5'-phospho polynucleotides in the presence of bivalent metal ions under neutral or alkaline conditions [1], occurs in many tissues of animals and plants and also in micro-organisms. Bovine pancreatic DNase I (EC 3.1.21.1) is the DNase I that has been studied most thoroughly [2]. Interestingly, cyanobacterial [3], bacterial [4], fungal and mitochondrial nucleases [5] also contain a DNase activity similar to that of bovine DNase I. However, the sequences of these nucleases are distinct from those of DNases I of bovine and other vertebrates [6–11]. To understand the evolutionary relationships of the bovine DNase-I-like proteins, we investigated DNase I in species of greater evolutionary separation: shrimp DNase I was therefore purified and characterized [12]. Although its enzymic properties are typical of DNase I, there were several distinctive structural features, such as a lack of glycosylation, high Cys content, high molecular mass and protease resistance [12,13]. These features prompted us to study its structural and functional relationships.

In the present study, we determined the protein and cDNA sequences for shrimp DNase I. Surprisingly, the sequences exhibited no similarity to those of the bovine DNase-I-like proteins. However, in the motif of residues 205–255, a number of identical residues were found when aligned with the conserved active sites of several nucleases. This finding has prompted us to investigate the intrinsic RNase activity in shrimp DNase I. In the presence of Mg²⁺ and Ca²⁺, shrimp DNase I did indeed exhibit RNase activity. Therefore shrimp DNase I, which was named incorrectly, is referred to as shrimp nuclease throughout the rest of this paper.

EXPERIMENTAL

Materials

Live shrimp (Penaeus japonicus) were obtained from the local market. Calf thymus DNA and dithiothreitol were purchased from Sigma. Shrimp nuclease was purified [12] and its DNase I activity was assayed as described previously [14,15]. Trypsin [treated with tosylphenylalaninychloromethane [‘TLCK’)] and chymotrypsin (TLCK-treated) were from Worthington. Iodoacetic acid and 2-mercaptoethanol were obtained from Wako. Iodo[2,14C]acetic acid (50 μCi/μmol) was from Amersham.

Protein sequencing

Determination of amino acid composition, peptide cleavage and separation were performed as described previously [7,16]. Except for disulphide pairing, the protein was reduced and S-carboxymethylated before digestion with protease [17]. Peptides were sequenced on an Applied Biosystems Sequencer, model 477A.

RNA isolation and reverse transcriptase-mediated PCR

Total RNA was isolated from 0.4 g of hepatopancreas of live shrimp by the guanidinium/phenol/chloroform method [18]. A cDNA library was synthesized from total RNA with oligo(dT) primers and Moloney-murine-leukaemia virus reverse trans-
criptase (Novagen). 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).

Rapid amplification of cDNA ends (RACE)

By using the cDNA library with a 5'-degenerate primer 5'-CCNCCAYGARCTAYTATTGC-3' (sense; shrimp nuclease residues 201–208) and a 3'-degenerate primer 5'-TCR- TTNACRTCAARCTDCTAC-3' (anti-sense; shrimp nuclease residues 342–349), a fragment of shrimp nuclease cDNA was obtained. On the basis of the nucleotide sequence of this cDNA fragment, specific primers were then synthesized for 3' and 5' RACE (Life Technologies). For 3' RACE, the primer 5'-GGAATCCTGACCTTGGATGACATCAATG-3' (sense; shrimp nuclease residues 272–280) was used. For 5' RACE, the primer 5'-ATCTCCACCGGTTGCATTGA-3' (anti-sense; shrimp nuclease residues 278–285) was used for first-strand synthesis and the nested primer 5'-GGAATTTCTAGTCCGCTTGTAACCT-3' (anti-sense; shrimp nuclease residues 249–256; underline indicates EcoRI cloning site) was used in PCR amplification.

Gene cloning and DNA sequencing

The PCR products of 5' and 3' RACE were purified with GeneClean III (BIO 101). The recovered DNA fragments were treated with Sall and EcoRI (Boehringer Mannheim) and ligated into pBS(+) plasmid (Stratagene). This inserted plasmid was transformed into Escherichia coli strain TG1. Recombinant plasmids were isolated and sequenced on a DNA sequencer (Perkin-Elmer, model 373).

Disulphide pairing

Intact shrimp nuclease (1 mg) was treated with CNBr in 70% (v/v) formic acid for 24 h at 25 °C [6] and the products were separated by gel filtration on a Sephadex G-75 column (0.9 cm × 67 cm) eluted with 50% (v/v) acetic acid. Two fragments were obtained and each was digested further with chymotrypsin (2 μg). The digests were separated on HPLC; Cys-containing peptides were identified. They were reduced with 2 mM dithiothreitol in 0.05 M Tris- HCl, pH 8.0, and chromatographed again on HPLC. Each separated into two reduced Cys-containing peptides, which were identified on the basis of results of amino acid composition and sequencing. The two reduced Cys-containing peptides were thus assigned for each of the paired disulphides.

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, 20 μl of 25 mM CaCl₂, 5 μl of 0.25 M MgCl₂, 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, then centrifuged. The A₅₆₀ of the supernatant was measured. One unit of RNase activity was defined as an increase of one absorbance unit under these conditions.

RNase activity staining in situ

The zymogram method was used as described by Rosenthal and Lacks [19], with minor modifications. Baker’s yeast RNA (Worthington) was added to a final concentration of 160 μg/ml in the separating gel solution. After electrophoresis, the gel was soaked in 100 ml of 10 mM Tris/HCl, pH 7.0, containing 5 mM MgCl₂ at 25 °C for 30 min with gentle shaking. This process was repeated twice. Finally, the gel was transferred to fresh buffer also containing 2 mM CaCl₂ at 37 °C for 30 min and stained with ethidium bromide. RNase activity was revealed as dark bands in the gel. SDS/PAGE was performed by the method of Laemmli [20].

RESULTS

Amino acid and cDNA sequences of shrimp nuclease

Approximately 95% of the amino acid sequence of shrimp nuclease was derived from protease-digestested peptides and agreed well with the cDNA-translated sequence (Figures 1 and 2). The uncertain parts of the protein sequence were supplemented by

Figure 1  Complete amino acid sequences of shrimp nuclease

Reduced and S-carboxymethylated nuclease was digested by trypsin. The large tryptic peptides were digested further by thermolysin or chymotrypsin. The intact nuclease cleaved with CNBr and its products were digested by chymotrypsin. Designations: T, tryptic peptide; Th, thermolytic peptide; C, chymotryptic peptide; CB, CNBr peptide. Daughter peptides are designated with letters after the names of the parental peptides. Solid lines, sequences completed by Edman degradation; broken lines, sequences by composition only; dotted lines, incomplete sequencing by Edman degradation.
Sequencing and cloning of shrimp nuclease

The nucleotide sequence was determined by the rapid amplification of cDNA ends (RACE) method using the SMARTER kit and the RT-PCR method. The translated amino acid sequence was deduced from the nucleotide sequence and is shown in Figure 2. The putative signal peptide, which is located at the N-terminus of the protein, is underlined.

Determination of the blocked N-terminus

It was not possible to determine the N-terminus of either the native or the denatured shrimp nuclease by Edman degradation on the sequencer. During the sequencing of the tryptic peptides, one of the peptides (T1) also failed to be sequenced. The amino acid composition and the cDNA sequence suggested that peptide T1 should be assigned at the N-terminus and that its calculated molecular mass was 997.25 Da. The actual molecular mass of peptide T1 measured by matrix-assisted laser desorption ionization–time-of-flight MS (MALDI–TOF MS) was 980.61 Da (Figure 3). Based on these results, cyclization of Gln to form pyroglutamate at the N-terminus of shrimp nuclease is a reasonable deduction.

Pairing of disulphide bonds

The shrimp nuclease sequence revealed eleven Cys residues. In 8 M urea without 2-mercaptoethanol, treatment of shrimp nuclease with iodoacetate showed no carboxymethyl Cys residues, but after reduction and S-carboxymethylation in 8 M urea, 11 carboxymethyl Cys residues were obtained (results not shown), indicating that all of the eleven Cys residues are involved in disulphide bonds. The first three disulphide bonds (Cys1-Cys3, Cys4-Cys6, and Cys7-Cys9) were paired, on the basis of identification of the Cys-containing chymotryptic peptides obtained from the large CNBr fragment (Figure 4). Pairing for the last four Cys residues (Cys10-Cys13) poses a rather difficult problem because no protease can digest the bond between Cys10 and Cys13. A peptide isolated from the chymotryptic digest of the CNBr fragment (CB5 in Figure 1) contained all these four Cys residues. The sequencing results for this chymotryptic peptide showed that the pairings would be either Cys10-Cys12; Cys11-Cys13 or Cys10-Cys13; Cys11-Cys12.

Except for five intramolecular disulphide bonds mentioned above, the eleventh Cys residue must form a disulphide with an unknown thiol compound. We treated the enzyme with 2-mercaptoethanol in a buffer without denaturing agent, followed by S-carboxymethylation with iodo[2-14C]acetate. It was still...
Figure 5 Identification of the 11th Cys residue and its cross-linked thiol compound

(A) HPLC profile of the radiolabelled peptide. Shrimp nuclease was reduced with 2-mercaptoethanol (9 mM) in 0.1 M Tris/HCl, pH 8.0, for 30 min and then alkylated with iodo[2-14C]acetate (10 mM). The labelled enzyme after digestion with trypsin for 2 h was subjected to gel filtration on a Sephadex G-25 column (0.4 cm × 70 cm). The radioactive fractions collected were digested further with thermolysin and the products were run through the same Sephadex G-25 column. The 14C-containing fractions were then analysed by HPLC. The arrow indicates the elution position for the 14C-containing peptide. (B, C) MALDI–TOF-MS of native shrimp nuclease (B) and the reduced and alkylated shrimp nuclease (C). The reduction and alkylation procedure was the same as that in (A) except that unlabelled iodoacetate was used. The difference between the two molecular masses (42091–41546 Da) plus the molecular mass of acetate minus 1 (for a proton) is 604 Da. Because the error of MALDI–TOF-MS is approx. 0.1%, the estimated molecular mass for the thiol compound attached to the native shrimp nuclease is between 500 and 700 Da. The polypeptide with ten Cys residues cross-linked, one Cys residue free and a blocked pyroglutamate at the N-terminus has a calculated molecular mass of 41425 Da.

Fully active but with 1 mol of 14C incorporated per mol of the enzyme. The labelled enzyme was then digested with proteases and subsequently analysed on HPLC (Figure 5A). The sequencing results showed a peptide, NLGScmCGDGGV (in which cmC represents carboxymethyl Cys), which indicated that Cys101 of shrimp nuclease was radiolabelled. On the basis of MALDI–TOF-MS analysis (Figures 5B and 5C), the thiol compound linking to the 11th Cys at position 101 has an estimated molecular mass between 500 and 700 Da.

Figure 6 Alignment of the conserved motifs with selected nucleases from different organisms and proposal of the structure of the shrimp nuclease active site

Upper panel: amino acid sequence alignment showing a conserved motif within a few nucleases. Numbers at the right relate to the last residue on each line. Consensus residues are shaded. Abbreviations: Shp, shrimp; Sem, Serratia marcescens; Hum, Homo sapiens; CaI, Caenorhabditis elegans; Cuc, Cunninghamella cucumeris; Syr, Syncephalastrum racemosum; Scp, Schizosaccharomyces pombe; Hep, Helicobacter pylori; Bob, Borrelia burgdorferi; Trb, Trypanosoma brucei; Ara, Arabananaa sp. Lower panel: the active site of Serratia nuclease, drawn by the MOSCRIPT program. The side chains of the highly conserved residues are shown by sticks. The residues of Serratia nuclease are illustrated and the parentheses indicate the corresponding residues in shrimp nuclease. The potential functions of these residues were proposed on the basis of the S. marcescens studies (see the text).
Shrimp nuclease in a solution containing 0.1 M MnCl$_2$ and 0.1 M iodoacetate at 47 °C was rapidly inactivated. The reaction kinetics was similar to that for bovine DNase I [25]. The inactivation of bovine DNase I by iodoacetate was due to the alkylation of a single His residue [25] at position 134 [6]. His$^{134}$ was considered to be at the active site; a catalytic mechanism involving His$^{134}$ was proposed on the basis of the X-ray structure [26]. For the identification of this His residue, shrimp nuclease was first modified with iodo[2-14C]acetate. A major radioactive peptide was then isolated in the tryptic digest of the labelled enzyme and His$^{134}$ was the radiolabelled residue. His$^{134}$ in shrimp nuclease, corresponding to His$^{180}$ in the *Serratia* nuclease, was the highly conserved histidine residue (Figure 6, upper panel).

**DISCUSSION**

Although bovine DNase I and shrimp nuclease have similar enzymic specific activities [12], their protein sequences are different. In the present study, sequence alignment revealed that the DNase I activity-containing proteins can be grouped into two categories. The first group is bovine DNase-I-like proteins, which is highly conserved and is found only in vertebrates [27]. The second group consists of some nucleases distributed widely in organisms ranging from cyanobacterium to human [3–5, 28]. The amino acid sequences for the second group are all different except in a 50-residue segment that is known to be the nuclease active site. Shrimp nuclease belongs to the second group. In evolution, shrimp has not yet quite evolved to differentiate its hepatopancreas. For digestive purposes it employs an ancient molecule, a sugar-non-specific nuclease, rather than the DNA-specific DNase I that is used in the more evolved vertebrates, which have differentiated liver from pancreas.

To the best of our knowledge, shrimp nuclease is the first nuclease to have been cloned and characterized among the Arthropoda. A sequence similarity search revealed no similar genes. Besides several distinctive properties in structure [12, 13], the present study demonstrated two unique features: a blocked N-terminus and an unusual mixed disulphide. These two features have never previously been reported among all known nucleases or DNases I. From the results of this study, we established not only the exact N-terminus and an odd number of Cys residues forming an unusual mixed disulphide. These two features have never previously been reported among all known nucleases or DNases I. From the results of this study, we established not only the exact number and locations of the Cys residues but also the pairing of disulphide bonds. Shrimp nuclease contained 11 Cys residues. They formed five intramolecular disulphide bonds; the eleventh Cys at position 101 was linked to a thiol compound with an estimated molecular mass of between 500 and 700 Da. The enzymes in which Cys$^{101}$ has been reduced or alkylated or disulphide-linked are all active, indicating that Cys$^{101}$ is not involved in catalysis. It has been reported [29–31] that a thiol group of a protease was cross-linked to thiomethanol. The enzyme activity could be generated only by reduction to remove the thiol compound. Recently an anti-tumour RNase from *Rana pipiens* oocytes, called onconase, was reported to have a pyroglutamate residue at the N-terminus [32]. A recombinant mutant onconase (N1M), with a Met residue replacing pyroglutamate, had decreased enzyme activity and cytotoxicity [33]. The biological significance of a blocked N-terminus and an unusual mixed disulphide bond in shrimp nuclease has yet to be elucidated.

The conserved active site found in the family of sugar non-specific nucleases (Figure 6) from organisms with distant evolutionary origins might imply that their catalytic mechanisms are similar. In particular, His$^{134}$ and Asn$^{145}$ in shrimp nuclease are also highly conserved in two other different families of nucleases.
the DNA-entry nuclease of the *Streptococcus pneumoniae* family and the Cys-His box containing the nuclear homing endonuclease family. The crystal structure of I-PpoI homing endonuclease complexed with a homing-site DNA oligonucleotide revealed that the conserved Asn residue ligates the essential bivalent cation, and the conserved His interacts with the negative charge on the scissile phosphate oxygen; it was suggested that the His serves as a Lewis acid to stabilize the transition state and that a metal-bound water molecule is deprotonated by a general base and attacks the phosphate group [34]. However, on the basis of mutational and biochemical studies, Friedhoff and co-workers [35] preferred the notion that the conserved His residue functions as a general base in the activation of the water molecule and the bivalent cation as a Lewis acid in phosphodiester bond cleavage.

In view of available structural, mutational and biochemical data of *Serratia* nuclease and I-PpoI, we predict that the active site structure of shrimp nuclease is similar to that of *Serratia* nuclease (Figure 6, bottom panel). In this structure, His$^{311}$ and Asn$^{314}$ are very important for catalytic activity, as mentioned above. Lys$^{316}$ and Arg$^{324}$ might be involved in substrate binding. The *Serratia* nuclease R87K mutant maintained wild-type activity, showing the importance of a positive charge at this position (Lys$^{316}$ in shrimp nuclease). Other conserved residues, including small residues, bulky hydrophobic residues, Gly$^{319}$, Asn$^{324}$, Pro$^{325}$, Gin$^{326}$, and Trp$^{328}$, may be involved in the structure stabilization of the active site. In particular, Asn$^{324}$ is involved in a hydrogen-bond network supporting the proper orientation of the imidazole ring of His$^{311}$ to abstract a proton from the water molecule, as well as Gin$^{326}$ to Asn$^{314}$, and Trp$^{328}$ to Glu$^{319}$. Moreover, because this nuclease’s active site is distinguishable at the sequence level, it can be used as a fingerprint for putative new members belonging to this subfamily.

We have expressed the cloned shrimp nuclease gene in *E. coli* but the expressed protein failed to show nuclease activity. This was probably because shrimp nuclease has more than five disulphide bonds to cross-link. The previous study [13] showed that once the disulphide bonds were reduced, the polypeptide chain was unable to fold back properly to regain its enzymic activity. The difficulties encountered prevented us from a further study of the important role of His$^{311}$ by site-directed mutagenesis. Nevertheless, the site-specific modification of His$^{311}$ has clearly demonstrated its important catalytic role.

This work was supported in part by a grant from the National Science Council of the Republic of China (NSC87-2316-B-002-010).

References


© 2000 Biochemical Society