Protein Science (1997), 6:2655–2658. Cambridge University Press. Printed in the USA.

FOR THE RECORD

Expression, characterization, and crystallization of a member of the novel phospholipase D family of phosphodiesterases

YI ZHAO, JEANNE A. STUCKEY, DANIEL L. LOHSE, AND JACK E. DIXON

Department of Biological Chemistry, The University of Michigan Medical School, Ann Arbor, Michigan 48109-0606

(RECEIVED August 12, 1997; ACCEPTED September 15, 1997)

Abstract: A family of phospholipase D (PLD) proteins has recently been identified (Koonin, 1996; Ponting & Kerr, 1996) based upon amino acid sequence identity. This family includes human and plant PLDs, proteins encoded by open reading frames in pathogenic viruses and bacteria, as well as an endonuclease. The endonuclease, known as Nuc, is encoded by the IncN plasmid, pKM101, present in *Salmonella typhimurium*. The recombinant Nuc protein has been expressed and purified from *Escherichia coli*. The amino-terminal sequencing of the purified protein indicated that the mature protein started from the 23rd residue of the predicted sequence, suggesting that the protein is proteolytically processed during export to the periplasmic space. The recombinant enzyme was able to hydrolyze both double and single-strand DNA and an artificial substrate, bis(4-nitrophenyl) phosphate, which contains a phosphodiester bond. The enzyme activity was not inhibited in the presence of EDTA and was not regulated by divalent cations. The purified protein has been crystallized by hanging drop vapor diffusion methods, and those crystals diffract to 1.9 Å resolution.

Keywords: crystallization; nuclease; phosphodiester; phospholipase D

Until recently phospholipase D (PLD) proteins were thought to be a restricted family of enzymes; however, with the emergence of new analysis programs, PLDs were found to have sequence identities to a variety of proteins, including two classes of phospholipid-specific enzymes (cardiolipin synthases and phosphatidylserine synthases), poxvirus envelope proteins, a murine toxin from *Yersinia pestis* (Cherepanov et al., 1991), a bacterial endonuclease (Nuc) (Pohlman et al., 1993), and an *E. coli* protein (YjhR) containing an N-terminal domain showing similarities with helicase (Burland et al., 1995). This sequence identity shared by this group of proteins indicates that they form a superfamily of enzymes (Koonin, 1996; Ponting & Kerr, 1996), that are capable of catalyzing the hydrolysis or formation of phosphodiester bonds. An alignment of the protein family demonstrates an invariant motif consisting of $HxK(x)_4D(x)_6G$ (Fig. 1A, B).

Interestingly nearly all the members of the superfamily have a duplication of the $HxK(x)_4D(x)_6G$ sequence motif, as is seen in the *Y. pestis* murine toxin and human PLD sequence shown in Figure 1A and B. The only known proteins that contain a single copy of the motif are a plasmid-encoded endonuclease Nuc (Pohlman et al., 1993) and the uncharacterized *E. coli* protein YjhR (Burland et al., 1995). Nuc is encoded by a 35.4 kilobase IncN plasmid pKM101 present in *Salmonella typhimurium* (Lackey et al., 1977).

There are no structures available for any member of the PLD superfamily, and there is also a limited understanding of the catalytic mechanism employed by this family of enzymes. This paper provides the first structural information for a member of the PLD family of proteins, and it is likely to serve as a model for the bacterial, viral, and human enzymes.

Results and discussion: The Nuc gene was inserted downstream of the T7 polymerase promoter and ribosomal binding site in the pT7-7 expression vector (Tabor & Richardson, 1985). The IPTG-induced recombinant protein was localized to the bacterial periplasmic space. One liter of bacterial culture produced approximately 10 mg of recombinant protein. After two chromatography steps (a Mono-S Sepharose ion exchange column and a Sephadex G-100 gel filtration column) the resulting protein was judged to be greater than 99% pure by Coomassie brilliant blue staining of a reducing SDS-PAGE gel (Fig. 1C, left panel). Sequencing of the amino terminus confirmed that the mature recombinant protein starts at residue 23 of the predicted protein sequence, suggesting that the signal peptide was efficiently cleaved when the protein was secreted into the periplasmic space.

The recombinant Nuc eluted from the gel filtration column as a monomer and is capable of hydrolyzing double-stranded DNA (Fig. 1C, middle panel) and single-stranded deoxy-oligonucleotides with an optimal pH range of 5.5–6.0 (data not shown). To obtain the catalytic constants of Nuc, a continuous spectrophotometric

Reprint requests to: Jack E. Dixon, Department of Biological Chemistry, The University of Michigan Medical School, 5416 Medical Science Building I, Ann Arbor, Michigan 48109-0606.

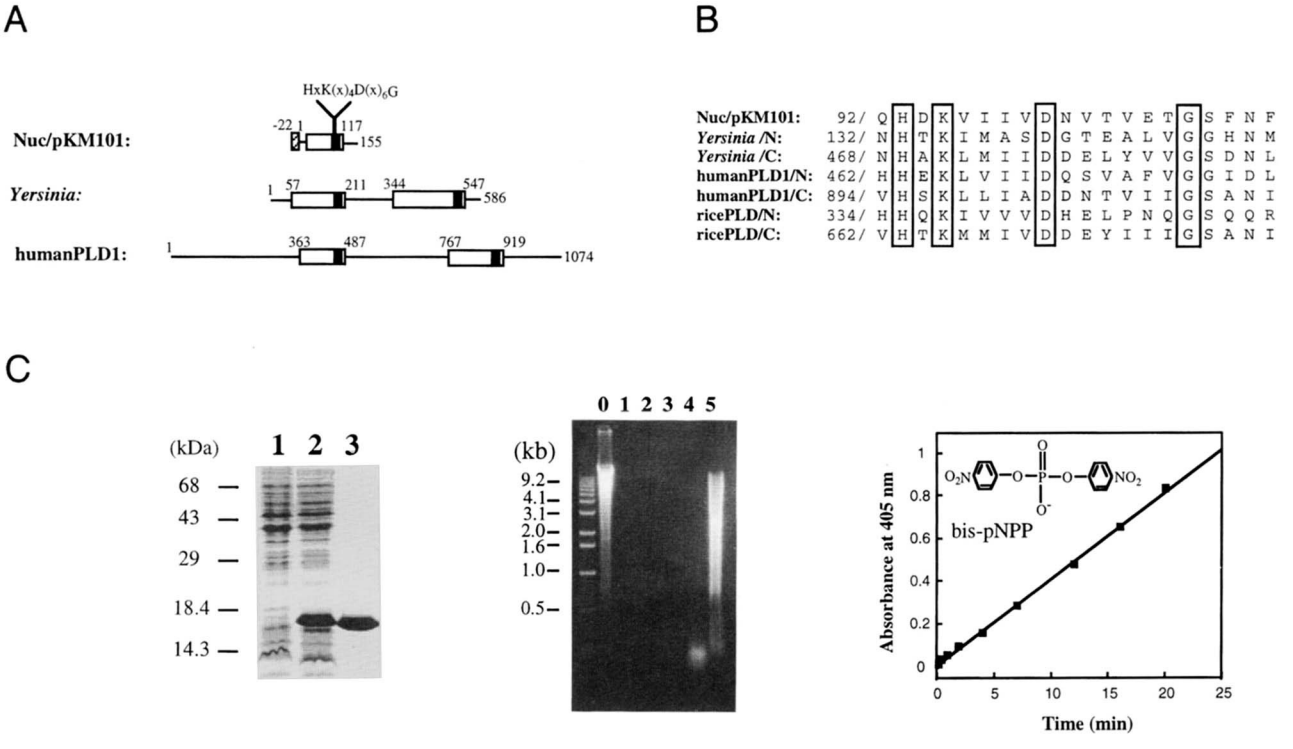

Fig. 1. Characterization of the Nuc gene product. (**A**) The schematic structure of Nuc/pKM101 (Pohlman et al., 1993), *Yersinia* murine toxin (Cherepanov et al., 1991), and human phospholipase D1 (PLD, Hammond et al., 1995). The open boxes represent PLD domains. The solid portions represent the highly conserved regions containing the $HxK(x)_4D(x)_6G$ motif sequence. The shaded box at the beginning of Nuc represents a signal sequence of 22 amino acids. The amino acids are numbered as shown. (**B**) Alignment of the amino acid sequences containing the $HxK(x)_4D(x)_6G$ motif in Nuc, *Yersinia* murine toxin, human PLD1, and rice PLD (Ueki et al., 1995). N and C PLD domains in *Yersinia* murine toxin and PLDs are designated. The highly conserved His, Lys, Asp, and Gly residues are boxed. The numbers at the beginning of each sequence indicate the number of amino acid residues from the N-terminus. (**C**) Expression and purification of Nuc (left panel). Cell extracts containing the pT7-7 vector only (lane 1) or with an inserted Nuc gene (lane 2) were separated on a 15% SDS-PAGE along with the purified Nuc product (lane 3). Protein standards are shown to the left of gel. Nuc activity is depicted in the middle and right panels. The middle panel shows nuclease activity using double-stranded calf thymus DNA as the substrate. Calf thymus DNA (7.5 μg) was incubated with varied amounts of Nuc for 10 min at 37 °C and then applied to a 1% agarose gel for separation. The DNA was visualized under UV light. 0 = no enzyme added as control; 1 = 4.0 μg; 2 = 0.8 μg; 3 = 0.16 μg; 4 = 0.032 μg; 5 = 0.0064 μg purified enzyme. The molecular size of DNA is shown on the left. Hydrolysis of bis(4-nitrophenyl) phosphate by Nuc is shown in the right panel. The hydrolyzed product, 4-nitrophenol, was detected by its absorbance at 405 nm. The final concentration of purified enzymes in the reaction mixtures was 0.017 mg/mL.

assay was developed using bis(4-nitrophenyl) phosphate. When the phosphodiester bond of bis(4-nitrophenyl) phosphate is cleaved, p-nitrophenol is released, and the increase in absorbance can be measured at 405 nm. Nuc hydrolyzes this substrate with a K_m of 8.5 mM and a k_{cat} of 0.12 s^{-1} at pH 5.9. Unlike most other nucleases, Nuc activity is not inhibited in the presence of EDTA (concentrations up to 1 mM) and is not influenced by the presence of divalent metals such as Mg^{2+}, Zn^{2+}, Ca^{2+}, and Mn^{2+} (data not shown).

Analysis of the X-ray structure of Nuc should provide the first three-dimensional information about the structure of any PLD family member as well as offer insights into the specific roles of the conserved Asp, His, and Lys residues (Fig. 1A). Large well-formed crystals of Nuc (Fig. 2) can be grown in eight days using the hanging drop method of crystallization. The crystals form in the tetragonal space group $P4_32_12$ with unit cell dimensions of a = 59.2 Å, b = 59.2 Å, c = 111.3 Å, and $\alpha = \beta = \gamma = 90.0°$. The crystals have a calculated molecular volume of 2.86 that constitutes a 55% solvent content when assuming one monomer per asymmetric unit. A 1.9 Å resolution native data set (94% complete in the 2.0–1.9 Å resolution shell) with an Rsym = 3.5% has been collected on a Rigaku RU-200 CuKα rotating anode X-ray generator equipped with a San Diego Multiwire dual detector system and processed with the accompanying software. Structure determination of the native protein is proceeding by incorporating multiple isomorphous replacement methods with multiple wavelength anomalous diffraction techniques (Hendrickson, 1991). Initial phases calculated from heavy atom derivatives have produced an interpretable electron density map of the protein, and tracing of the protein backbone is in progress. The structure of Nuc will serve as the architectural basis for the catalytic domain of the superfamily of phospholipases. The structural determinations of the native enzyme and catalytically inactive mutants complexed with substrate should suggest roles for the key residues involved in catalysis and serve as a mechanistic paradigm for this family of phosphodiesterases.

Materials and methods: *Expression and purification of the plasmid-encoded endonuclease Nuc:* The coding region of Nuc

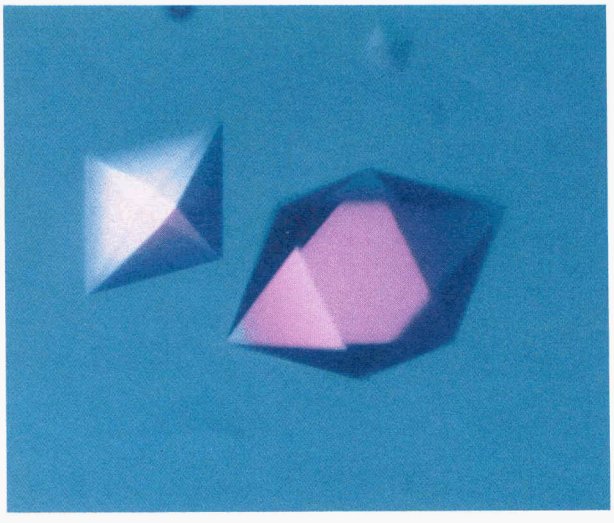

Fig. 2. Crystals of Nuc. Tetragonal crystals of Nuc exhibiting a characteristic bipyramidal shape. Crystals are shown at 75× magnification under polarized light that produces the color variations. Actual size of the largest crystal is 0.4 × 0.4 × 0.6 mm^2.

was amplified by polymerase chain reaction (PCR) from the bacterial plasmid pKM101. PCR primers introduced *Nde*1 (with the initiation codon ATG) and *Bam* H1 restriction sites at the beginning and the end of the coding region. The PCR amplified fragment was inserted into *Nde* 1/*Bam* H1 sites of the bacterial expression vector pT7-7 and the resultant construct transformed into *E. coli* BL21(DE3) cells. One liter of 2xYT media containing 100 μg/mL ampicillin was inoculated with a freshly grown overnight colony, and the culture was grown at 37 °C to an $O.D._{600}$ of 0.6–0.7. Expression of the recombinant Nuc was induced with 0.4 mM isopropyl-thio-β-galactosidase (IPTG) for four hours at room temperature. The harvested cells were resuspended in 20 mL of cold buffer containing 20% sucrose, 30 mM Tris-HCl (pH 8.0), and 1 mM EDTA, incubated on ice for 10 min, then centrifuged at 10,000 × *g*. The supernatant (I) was collected and the bacterial pellet was resuspended in 20 mL of 5 mM $MgSO_4$ (cold) for 10 min, and then the suspension was centrifuged at 10,000 × *g* for 10 min. The supernatant (II) was collected and combined with the supernatant (I).

For protein purification, the supernatants (I and II) were applied to a Mono-S Sepharose (Sigma) column that was pre-equilibrated with 30 mM Tris-HCl (pH 8.0). The column was extensively washed with buffer and the bound proteins were eluted with a linear gradient of 0 to 1 M NaCl. The collected fractions containing Nuc were pooled, and the pH was adjusted to 7.5. The protein was concentrated to 10 mg/mL, and then applied to a Sephadex G-100 gel permeation chromatography column equilibrated with 30 mM Tris-HCl (pH 7.5) and 150 mM NaCl.

Enzymatic activity assay: For the endonuclease assay, calf thymus DNA (Sigma) was used as the substrate. Ten milligrams of the purchased calf thymus DNA was dissolved in 10 mL of 10 mM Tris-HCl (pH 8.0) and 5 mM EDTA, extracted six times with phenol/chloroform, and then precipitated with ethanol. The DNA pellet was extensively washed with 70% ethanol and dissolved in water to a final concentration of 1.5 mg/mL. Endonuclease activity of the purified Nuc was assayed by incubating varied amounts of enzyme with 7.5 μg of calf thymus DNA in 50 mM Tris-HCl, 50 mM Bis-Tris, and 100 mM acetate (pH 5.9) for 10 min at 37 °C in a final volume of 20 μL. The resulting assay mixtures were run on a 1% agarose gel and stained with ethidium bromide to determine the extent of DNA cleavage. The artificial substrate bis(4-nitrophenyl) phosphate was also used for the activity assays. The reactions were performed in 50 mM Tris-HCl, 50 mM Bis-Tris, and 100 mM acetate (pH 5.9) with 25 mM bis(4-nitrophenyl) phosphate at 37 °C in a volume of 0.3 mL, then stopped by addition of 0.7 mL of 0.25 M sodium hydroxide. The enzyme activity was measured as an increase in absorbance at 405 nm (Zhang & Van Etten, 1991).

Crystallization of Nuc: A purified sample of Nuc in 150 mM NaCl, 30 mM Tris-HCl (pH 7.5) was concentrated to 10 mg/mL and buffer exchanged to 10 mM NaCl, 5 mM Tris-HCl (pH 7.5) using a Centriprep-10 (Amicon). Initial crystallization conditions were determined using the sparse matrix sampling kit (Hampton Research, La Jolla, CA) (Jancarik & Kim, 1991) employing the vapor diffusion hanging drop technique. The best crystals were grown from drops containing 5 μL of a 10 mg/mL protein solution mixed with 5 μL of the precipitating solution (2 M ammonium sulfate, 100 mM Tris-HCl, pH 7.5). The crystals nucleated in four days at room temperature (~22 °C) and grew to a size of 0.4 × 0.4 × 0.6 mm^3 within eight days.

Diffraction analysis of Nuc crystals: Nuc crystals were characterized on a Rigaku RU-200 rotating anode equipped with the San Diego multiwire dual detector system. Before mounting, the crystals were soaked for five minutes in a stabilizing solution containing 5% glycerol, 2 M ammonium sulfate, 100 mM Tris-HCl (pH 7.5) then soaked for 30 to 60 min in a stabilizing solution containing 20% glycerol. The crystals were mounted in a 0.3 mm loop and flash frozen to −150 °C in a liquid nitrogen stream to prolong the life of the crystals during data collection.

Acknowledgments: We thank Dr. S.C. Winans for providing the plasmid pKM101, and Dr. Elizabeth Gottlin for discussions and critical reading of this manuscript. This work was supported by a grant from the National Institutes of Health, NIDDKD 18849 to J.E.D. and by the Walther Cancer Institute. Y.Z. is supported by an NIH (NIDDKD) National Research Service Award postdoctoral fellowship.

References

Burland V, Plunkett GI, Sofia HJ, Daniels DL, Blattner FR. 1995. Analysis of the *Escherichia coli* genome VI: DNA sequence of the region from 92.8 through 100 minutes. *Nucleic Acids Res 23*:2105–2119.

Cherepanov PA, Mikhailova TG, Karimova GA, Zakharova NM, Ershov YV, Volkovoi KI. 1991. Cloning and detailed mapping of the fra-ymt region of the Yersinia pestis pFra plasmid. *Mol Gen Mikrobiol Virusol 12*:19–26.

Hammond SM, Altshuller YM, Sung TC, Rudge SA, Rose K, Engebrecht J, Morris AJ, Frohman MA. 1995. Human ADP-ribosylation factor-activated phosphatidylcholine-specific phospholipase D defines a new and highly conserved gene family. *J Biol Chem 270*:29640–29643.

Hendrickson WA. 1991. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science 254*:51–58.

Jancarik J, Kim SH. 1991. Sparse matrix sampling: A screening method for crystallization of proteins. *J Appl Crystallogr 24*:409–411.

Koonin EV. 1996. A duplicated catalytic motif in a new superfamily of phosphohydrolases and phospholipid synthases that includes poxvirus envelope proteins. *Trends Biochem Sci 21*:242–243.

Lackey D, Walker GC, Keng T, Linn S. 1977. Characterization of an endonuclease associated with the drug resistance plasmid pKM101. *J Bacteriol 131*:583–588.

Pohlman RF, Liu F, Wang L, More MI, Winans SC. 1993. Genetic and biochemical analysis of an endonuclease encoded by the IncN plasmid pKM101. *Nucleic Acids Res 21*:4867–4872.

Ponting CP, Kerr ID. 1996. A novel family of phospholipase D homologues that includes phospholipid synthases and putative endonucleases: identification of duplicated repeats and potential active site residues. *Protein Sci 5*:914–922.

Tabor S, Richardson CC. 1985. A bacteriophage T7 RNA polymerase/promotersystem for controlled exclusive expression of specific genes. *Proc Natl Acad Sci USA 82*:1074–1078.

Ueki J, Morioka S, Komari T, Kumashiro T. 1995. Purification and characterization of phospholipase D (PLD) from rice (*Oryza sativa* L.) and cloning of cDNA for PLD from rice and maize (*Zea mays* L.). *Plant Cell Physiol 36*:903–914.

Zhang ZY, Van Etten RL. 1991. Pre-steady-state and steady-state kinetic analysis of the low molecular weight phosphotyrosyl protein phosphatase from bovine heart. *J Biol Chem 266*:1516–1525.