Vol. 203, No. 1, 1994 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 30, 1994 Pages 182-189

THE COMPLETE PRIMARY STRUCTURE ELUCIDATION OF *ASPERGILLUS FICUUM (NIGER)*, pH 6.0, OPTIMUM ACID PHOSPHATASE BY EDMAN DEGRADATION

Abul H.J. Ullah, Edward M. Mullaney, and H. Charles Dischinger, Jr.

Southern Regional Research Center, ARS, USDA, New Orleans, LA 70124

Received June 17, 1994

SUMMARY: The primary structure of the *Aspergillus ficuum* (*niger*) NRRL 3135 extracellular, pH 6.0,optimum acid phosphatase (E.C.3.1.3.2) was elucidated by gas phase sequencing. It was deduced by sequence overlap of peptides obtained from trypsin, chymotrypsin, clostripain, and cyanogen bromide digests of the pyridylethylated protein. The mature, active protein is composed of 583 amino acids, including 13 glycosylated Asn residues. The unglycosylated protein has a MW of 64,245-KDa and a pI of 4.97. Two putative metal binding sites were identified in the molecule. This enzyme may represent a special class of high molecular weight acid phosphatase, since it lacks the active site sequence RHGXRXP and shows no significant homology with known acid phosphatases containing this active site. Homology to human type 5 and *A.niger* APases was detected, however.

Acid phosphatases (APases, orthophosphoric monoester phosphohydrolase, E.C.3.1.3.2) are ubiquitous enzymes capable of hydrolyzing a wide variety of phosphomonoesters (1). Three classes of acid phosphatases can be distinguished on the basis of pH optima: highly acidic APases, with optima between pH 2.5 and 2.2; somewhat acidic, pH 5.5 - 5.0; and mildly acidic, pH 6.0 (2,3,4). Some of the members of this class of enzymes degrade *myo*-inositol hexakisphosphate, phytic acid, in a stepwise manner to lower levels of phosphorylation (5,6). *Aspergillus ficuum* (*niger*) NRRL 3135 when grown under limited phosphate in starch medium, produces three extracellular acid phosphatases. Two, with pH optima of 5.5 and 2.5 (E.C.3.1.3.8 and 3.1.3.2, respectively), are phytases (2,3). The third enzyme is an APase (E.C. 3.1.3.2) with a pH optimum of 6.0 (4). When purified and concentrated to greater than 1.0 mM, the pH 6.0 optimum APase has an intense blue color with an absorption at 580 nm (4). The enzyme lacks phytase activity, is thermostable at 63°C, has a weakly acidic pH optimum, and is a metalloenzyme. These last three features might be of commercial importance. We hope to engineer the substrate-binding domain to enable the enzyme to hydrolyze phytate. We report here the complete covalent structure of the active pH 6.0 optimum APase as deduced by chemical sequencing.

0006-291X/94 $5.00

MATERIALS AND METHODS

Protein purification: *Aspergillus ficuum* extracelluar pH 6.0 optimum APase was purified by sequential ion-exchange chromatography and chromatofocusing (4). A typical purification gave a specific activity of 3300 nKat/mg.

Reduction, alkylation, and repurification: Purified APase was derivatized using 4-vinylpyridine in the presence of DTT to modify cysteine (7). The pyridylethylated (PE) APase was repurified by reverse-phase on a SynChropak RP-8 column (0.46 x 25 cm, SynChrom, Lafayette, IN). The identity of the pH 6.0 optimum APase was verified by N-terminal sequencing. The purified pyridylethylated protein was used throughout the study for peptide mapping and sequencing.

Chemical and enzymatic cleavage of PE-acid phosphatase: PE-APase (1-2 nmole) was cleaved with cyanogen bromide (CNBr), trypsin, chymotrypsin, and clostripain (E.C. 4.4.22.8), as outlined by Tarr (8).

Peptide mapping and purification: Peptide maps of the CNBr, tryptic, and chymotryptic digests were developed using a standard bore (0.46 cm) C-18 (Vydac). For the clostripain digest, the map was established on narrow bore (0.21 cm) C-18 (Vydac). Elution gradients of aqueous 0.1% trifluoroacetic acid (TFA) against 0.1% TFA in acetonitrile were used.

Peptide sequencing: The primary structures of C-18 HPLC purified peptides were determined using a Porton PI2090 gas phase sequencer (Beckman Instruments, Fullerton, CA).

Sequence alignment and secondary structure prediction: Fungal pH 6.0 APase sequence was compared with the National Biomedical Research Foundation's Atlas protein sequence database (Release 37) using the FASTA program of Pearson and Lipman (9). Global and local alignments of fungal pH 6.0 optimum APase and human APase type 5 were performed by the ALIGN and LALIGN software provided by William Pearson (Univ. of Virginia, Charlottesville, VA). Secondary structure prediction for the protein was performed by the method of Garnier, Osguthorpe, and Robson (10).

RESULTS

Purification of pyridylethylated APase: The reduced and alkylated pH 6.0 optimum APase was eluted as a single component in C-8 HPLC at 55% (v/v) acetonitrile (data not shown). It had a single N-terminal sequence of **VVDERFPTYGPAVPIGDWVD**, which is in agreement with the published N-terminal sequence for fungal pH 6.0 optimum APase (4).

Cyanogen bromide fragmentation of PE-APase: Nine fragments were obtained upon treatment of the PE-Apase with cyanogen bromide (Figure 1). The cleavage products were eluted in a narrow range of the acetonitrile gradient (40-48%).

Tryptic map of PE-APase: Chromatography of the tryptic cleavage products of PE-APase resulted in 22 peaks (Figure 2). Some of the broad peaks contained more than one peptide.

Chymotryptic map of PE-APase: The three-step gradient used to separate peptides from the chymotryptic cleavage of APase produced 46 peaks (Figure 3). As in the tryptic digest, this procedure resulted in broad peaks which consisted of multiple peptides.

Clostripain map of PE-APase: The narrow-bore C-18 purification of peptides generated by treatment of APase by clostripain resulted in 16 peaks (Figure 4). This result is in agreement with the number expected based on the Arg content of the APase molecule. Peak 1 gave the sequence of the N-terminus, and peak 10 was identified as the C-terminus of the protein.

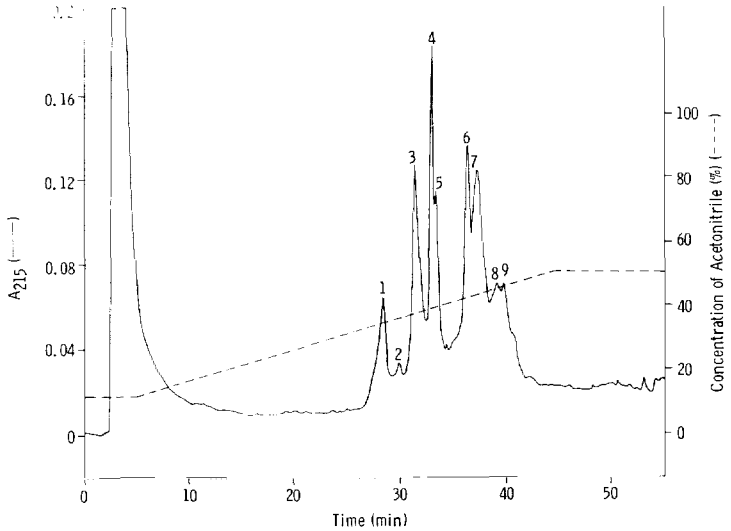

Figure 1. Elution profile of reverse-phase C18 chromatography of cyanogen bromide cleaved peptides of PE-APase.

Primary structure: The total covalent structure of fungal pH 6.0 optimum APase as elucidated by chemical sequencing is in Figure 5. The mature functional enzyme is a 583 amino acid protein with extensive glycosylation. The protein contains 6 cysteinyl residues, all of which are located in a stretch of 220 amino acids (residues 104 to 324).

Secondary structure: The secondary structure of fungal pH 6.0 optimum APase was predicted according to the method of Garnier, Osguthorpe, and Robson (10). The molecule has an

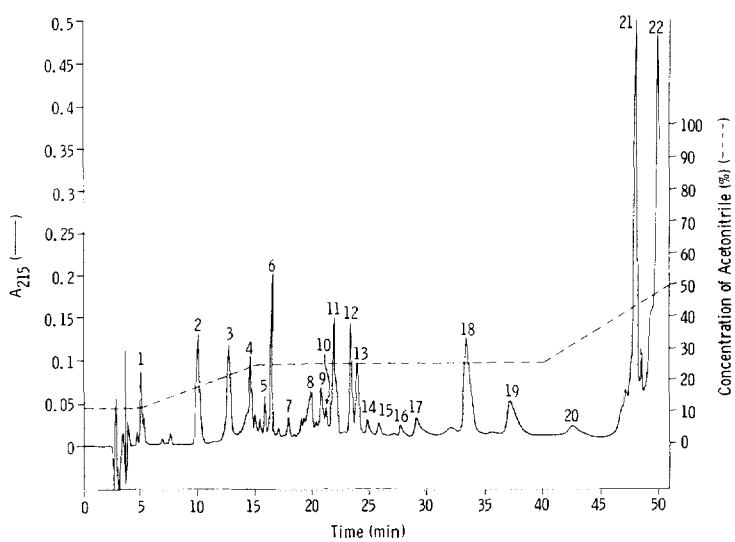

Figure 2. Elution profile of reverse-phase C18 chromatography of tryptic digests of PE-APase.

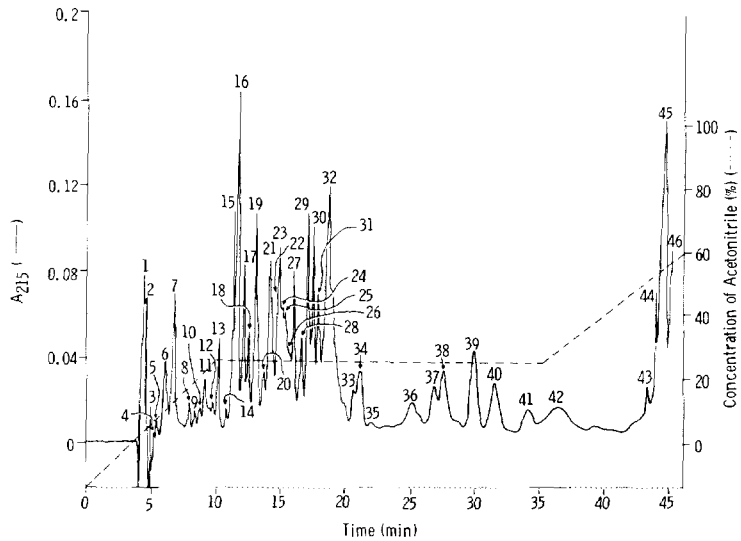

Figure 3. Elution profile of reverse-phase C18 chromatography of chymotryptic digests of PE-APase.

abundance of β-sheets (27 stretches) and very little α-helix (6 stretches) comprising 27.8%, and 11.8%, respectively. The β-turn and coil content of the protein is about 29.15% and 31.2%, respectively.

Location of Cys residues: Secondary structure prediction indicated 5 of the 6 cysteines are likely to be in β-turns; only Cys 291 is expected to be in an α-helix. The Kyte and Doolittle hydropathy plot (11) indicates that Cys 104, Cys 218, and Cys 324 are in hydophilic environments, while Cys 113, Cys 211, and Cys 316 are predicted to be in hydrophobic pockets.

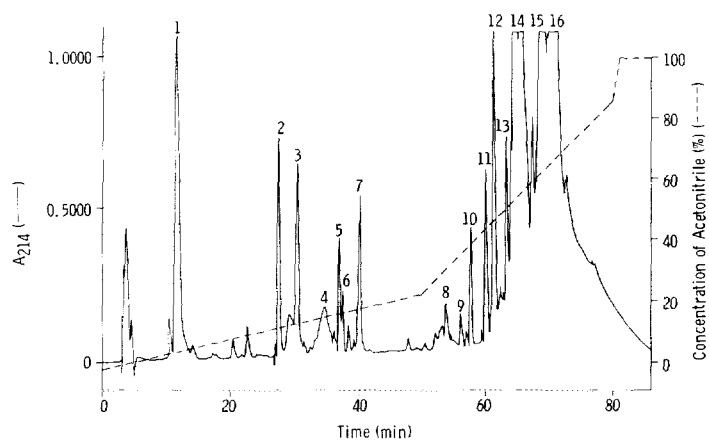

Figure 4. Elution profile of reverse-phase C18 chromatography of clostripain digests of PE-APase.

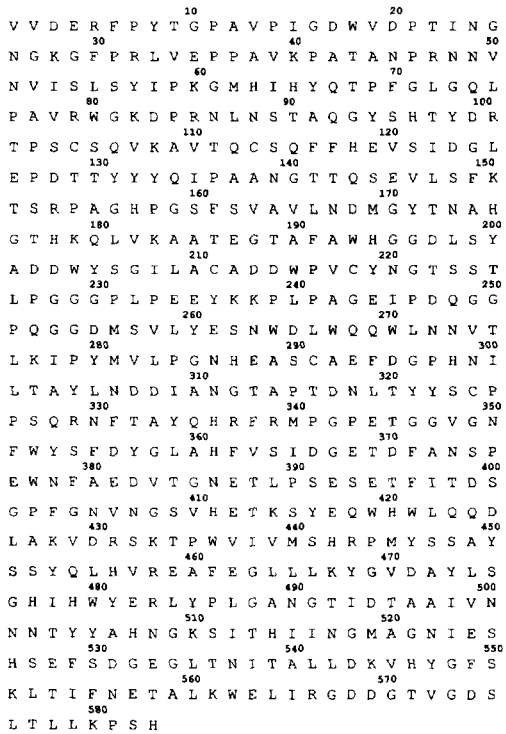

Figure 5. Primary structure of *Aspergillus ficuum* (*niger*), pH 6.0, optimum APase.

Modified amino acids: Native pH 6.0 optimum APase is extensively glycosylated. Sugar analysis detected only N-linked high mannose oligosaccharides (4). Thirteen blank cycles appeared during sequencing of the native molecule. All of these were followed by Thr or Ser after an intervening residue, indicating Asn-linked glycosylation. The other 23 Asn residues lacked glycosylation signals, and hence were left unmodified.

Molecular weight: The MW of the unglycosylated protein is calculated to be 64.24-KDa.

Isoelectric point: The pI of the protein was computed to be 4.97. The chromatographic behavior of the native enzyme is consistent with this value (4). The emperically-obtained pI of the native APase, 4.9, is essentially the same as that deduced from the primary structure.

Amino acid composition: The mature native pH 6.0 optimum APase is a 583 residue acidic protein. The protein consists of 38.75% non-polar, 40.49% polar, 10.8% acidic, and 9.95% basic residues. The Trp and His content of the protein is 15 and 23, respectively.

Metal binding site: The primary structure of pH 6.0 APase (Figure 5) includes two putative copper binding sites of the type described by Adman for ascorbate oxidase (12). The first is located in the N-terminal region and has the sequence **HIH** (residues 63-65). The second site has the same sequence but is in the C-terminus (residues 477-479).

Sequence similarities with other acid phosphatases: The FASTA (9) search of the NBRF protein databases indicated that a 140 amino acid region from *Aspergillus niger* acid phosphatase (16) shows 77.9% identity to *A. ficuum* pH 6.0 optimum APase. Another significant similarity was found in a 173 amino acid stretch of APase 5 precursor from rat (13). Local alignment of fungal pH 6.0 optimum APase with human type 5 APase revealed two regions of homology (Figure 6). A metal binding site with the sequence **H*H** was preserved in the larger segment. The smaller region (12 residues) was 58.3% homologous. While the first eight residues of this latter segment are consistent with β-sheet formation, the four C-terminal residues are in the coil conformation. A preserved His and an acidic residue reside in this region.

DISCUSSION

We have reported native *Aspergillus ficuum* pH 6.0 optimum APase to be an 82-KDa extracellular protein, with 28% glycosylation (4). Comparison of the calculated MW (from sequence data) with the published MW indicates 21.7% of the molecular weight is due to glycosylation. The sequence-derived molecular mass of fungal APase is higher than that of the two phytases from the same organism (2,3). Based on secondary structure prediction and hydropathy plot of the *Aspergillus ficuum* pH 5.0 phytase and pH 6.0 optimum APase, it is clear that these phosphomonoesterases are structurally unrelated proteins. While the phytases contain the active site residues **RHGXRXP**, APase lacks this sequence. Furthermore, for the pH 5.0 phytase, parallel α/β domains are expected, making it structurally similar to

```
                330       340       350       360       370
AFAPAS YYSCPPSQRNFTAYQHRFRMPGPETGGVGNFWYSFDYGLAHFVSIDGETDFANSPE-WNF
       .:     . : . .:. :    .. .     :: .. . .:. .......... .. :::
HUMACP FYFTGVQDINDKRFQETFEDVFSDRSLRKVPWYVLAGNHDHLGNVSAQIAYSKISKRWNF
               80        90       100       110       120       130

     380                  390       400       410       420
AFAPAS A-----------EDVTGNETLPSESETFITDSGPFGNVNGSVHETKSYEQWHWLQQDLAK
       .           .. .. ...  .. :. ..:. : . . . .  .. .:  ::...::
HUMACP PSPFYRLHFKIPQTNVSVAIFMLDTVTLCGNSDDFLSQQPERPRLTARTQLSWLKKQLAA
            140       150       160       170       180       190

      430       440       450       460       470
AFAPAS VDRSKTPWVIVMSHRPMYSSAYSSYQLHVREAFEGLLLKYGVDAYLSGHIH
          ... .:.: .: :..: :  .    . ...  :: .:::.:::.:: :
HUMACP ---AREDYVLVAGHYPVWSIAEHGPTHCLVKQLRPLLATYGVTAYLCGHDH
               200       210       220       230       240

        280
AFAPAS KIPYMVLPGNHE
       :.:. ::.:::.
HUMACP KVPWYVLAGNHD
               110
```

Figure 6. Alignment of the fungal,pH 6.0,optimum APase (AFAPAS) with the human type 5 APase (HUMACP) protein sequence. Alignment was accomplished by the method of Pearson and Lipman (9) using the program LALIGN. Sequence homology is indicated by a colon while conservative substitutions are identified by a period.

phosphoglycerate mutase (14) and rat acid phosphatase (15), which also have the **RHG** active site motif. For pH 6.0 optimum APase, however, we propose an antiparallel β-domain structure. Although the protein is not homologous with APases containing the **RHGXRXP** active site sequence, homology was found to one APase from *Aspergillus niger*, and human type 5 APase.

The optimum temperature for catalysis for pH 6.0 optimum APase is 63°C, five degrees higher than that of the pH 5.0 phytase (3). This enhanced thermal stability of pH 6.0 optimum APase may be due to the metal present in the structure of the protein. Both the blue color visible at 1.0 mM concentration and an absorbance peak at 580 nm (4) imply the presence of copper in the molecule. This is further supported by the presence of two sites with sequences which match the **H*H** motif for binding copper (12). Site-directed mutagenesis of these sites can reveal the role of metal in active site formation and catalysis.

Aspergillus niger, a close relative of strain NRRL 3135, produces an APase under similar conditions which has an homologous N-terminus and an internal 140 residue overlap (16). The clone of the *A. niger* gene was obtained and has been resequenced. The resultant corrected protein sequence is 96.5% homologous with the *A. ficuum* NRRL 3135 pH 6.0 APase (data not shown). A similar acid phosphatase with a pI of 4.1 and optimum pH of 6.1 for catalysis has been reported in *Aspergillus nidulans* (17). It is plausible that these three fungal APases are structurally and functionally related.

Future study of this metalloenzyme will be directed toward a more thorough understanding of structure/function relationships, and especially in reference to the phytases. Chemical probing of the active site will reveal whether a crucial His or Cys is involved in the formation of a phosphoenzyme transition complex, and whether Arg residues are involved in the substrate binding and anchoring, as predicted for *Aspergillus ficuum* (*niger*) phytases (18). In the absence of a crystallographically-derived three dimensional structure, model building by computer-aided molecular design will necessarily use distance geometry or homology modeling.

REFERENCES

1. Hollander, V.P. (1970) *In* The Enzymes (P.D. Boyer, Ed), pp. 449-498, Academic Press, New York, NY.
2. Ullah, A.H.J. and Cummins, B.J. (1987) Prep. Biochem. **17**:397-422.
3. Ullah, A.H.J. and Gibson, D.M. (1987) Prep. Biochem. **17**:63-91.
4. Ullah, A.H.J. and Cummins, B.J. (1988) Prep. Biochem. **18**:37-65.
5. Ullah, A.H.J. and Phillippy, B.Q. (1988) Prep. Biochem. **18**:483-489.
6. Ullah, A.H.J. and Phillippy, B.Q. (1994). J. Agric. Food Chem. **42**:423-425.
7. Thomsen, J. and Bayne, S. (1988) J. Protein Chem. **7**:295-296.
8. Tarr, G. (1986) *In* Methods of protein microcharacterization (J. Shively, Ed.), pp. 155-194, Humana Press, Clifton, NJ.
9. Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. **85**:2444-2448.
10. Garnier, J., Osguthorpe, D.J., and Robson, B. (1978) J. Mol. Biol. **120**:97-120.

11. Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. **157**:105-132.
12. Adman, E.T. (1991) *In* Advances in Protein Chemistry, Vol. 42 (C.B. Anfinsen, J.T. Edsall, F.M. Richards, and D.S. Eisenberg, Eds), pp. 145-197, Academic Press, San Diego, CA.
13. Ketcham, C.M., Roberts, R.M., Simmen, R.C., and Nick, H.S. (1989) J. Biol. Chem. **264**:557-563.
14. Richardson, J.S. (1981) *In* Advances in Protein Chemistry, Vol. 34 (C.B. Anfinsen, J.T. Edsall, and F.M. Richards, Eds.), pp. 167-339, Academic Press, New York, NY.
15. Lindqvist, Y., Schneider, G., and Vihko, P. (1993) J. Biol. Chem. **268**:20744-20746.
16. MacRae, W.D., Buxton, F.P., Sibley, S., Gwynne, R.W. and Arst Jr., H.N. (1988) Gene **71**:339-348.
17. Harsanyi, Z. and Dorn, G.L. (1972) J. Bacteriol. **110**:246-255.
18. Ullah, A.H.J., Cummins, B.J., and Dischinger, Jr., H.C. (1991) Biochem. Biophys. Res. Commun. **178**:45-53.