

UvA-DARE (Digital Academic Repository)

Primary structure, characterization and expression of the vanadium chloroperoxidase from the fungus Curvularia tnaequalis

Simons, B.H.; Barnett, P.; Vollenbroek, E.G.M.; Dekker, H.L.; Muijsers, A.O.; Messerschmidt, A.; Wever, R. **DOI**

10.1111/j.1432-1033.1995.0566k.x

Publication date 1995

Published in European Journal of Biochemistry

Link to publication

Citation for published version (APA):

Simons, B. H., Barnett, P., Vollenbroek, E. G. M., Dekker, H. L., Muijsers, A. O., Messerschmidt, A., & Wever, R. (1995). Primary structure, characterization and expression of the vanadium chloroperoxidase from the fungus Curvularia tnaequalis. *European Journal of Biochemistry*, *229*, 566-574. https://doi.org/10.1111/j.1432-1033.1995.0566k.x

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Primary structure and characterization of the vanadium chloroperoxidase from the fungus *Curvularia inaequalis*

Bert H. SIMONS¹, Philip BARNETT¹, Esther G. M. VOLLENBROEK¹, Henk L. DEKKER¹, Anton O. MUIJSERS², Albrecht MESSERSCHMIDT³ and Ron WEVER¹

E. C. Slater Institute, Department of Biochemistry, Plantage Muidergracht, Amsterdam, The Netherlands

² E. C. Slater Institute, Department of Biochemistry, Meibergdreef 15, Amsterdam, The Netherlands

3 Max-Planck-Institut für Biochemie, Martinsried, Germany

(Received 28 October 1994/23 January 1995) - EJB 94 1653/2

Using reverse transcription of messenger RNA followed by amplification using the polymerase chain reaction, three overlapping cDNA fragments encompassing the encoding sequence of the vanadium chloroperoxidase from the fungus Curvularia inaequalis were isolated and sequenced. The sequence was confirmed by DNA sequence analysis of genomic DNA. The deduced amino acid sequence predicts a protein of 609 residues with a mass of 67488 Da. Competitive reverse-transcription polymerase chain reaction analysis indicates that vanadium chloroperoxidase expression takes place in the secondary-growth phase initiated by nutrient depletion. Southern-blot analysis of genomic DNA indicates that there is only a single gene encoding the vanadium chloroperoxidase and that no isoenzymes are present. The Nterminal amino acid residue was blocked and could not be determined by amino acid sequencing, although evidence is presented showing that the N-terminal region starts very close to the first encoded methionine residue. Although the vanadium chloroperoxidase is secreted, it was not possible to assign a leader peptide. The protein contains two putative N-glycosylation sites but experiments indicate that the protein is non-glycosylated. Two cysteine residues are present in the protein both as free thiols; no disulphide bridging was found. Metal analysis revealed that iron, copper, and calcium do not constitute part of the protein. Zinc was found at a ratio of 0.3 ± 0.04 mol/mol protein. Boiling and subsequent SDS/PAGE of the protein sample showed a typical degradation pattern of the enzyme. Amino acid sequence analysis of the resulting peptides showed that the cleavage took place at Asp-Pro bonds of which six are located throughout the protein. No sequence similarity with other known peroxidases was found except for one small region, sharing limited similarity with bacterial haloperoxidases and other α/β -hydrolase-fold enzymes. In the case of the bacterial bromoperoxidases from this group, a methionine located in this region was suggested to have a role in catalysis. Methionine, however, was not involved in the catalysis of the vanadium chloroperoxidase.

Keywords. Vanadium chloroperoxidase; Curvularia inaequalis; cDNA; primary structure.

Haloperoxidases have been isolated from a broad range of organisms. These enzymes catalyze the oxidation of halides (X^-) in the presence of hydrogen peroxide to form hypohalous acids (HOX), which in turn can react with a nucleophilic acceptor [AH; Eqn (1), Eqn (3) and Eqn (4)] if present, to form a halogenated compound (AX) [1]. Specific halogenation [Eqn (1) and Eqn (2)] may, however, occur via a halogenating enzyme intermediate [2].

$$H_2O_2 + X^- + H^+ + Enzyme \rightarrow Enzyme \cdot HOX + H_2O_1$$
 (1)

$$Enzyme \cdot HOX + AH \rightarrow Enzyme + AX + H_2O$$
, (2)

Enzyme. Peroxidases (EC 1.11.1.-).

$$Enzyme \cdot HOX \rightarrow Enzyme + HOX,$$
 (3)

$$HOX + AH \rightarrow AX + H_2O.$$
 (4)

The haloperoxidases are named according to the oxidation of the most electrophilic halide they are able to catalyze, e.g. bromoperoxidases are able to oxidize iodide and bromide while chloroperoxidases are also able to oxidize chloride. Many naturally occurring halogenated compounds have been isolated and, as far as determined, they are involved in a variety of metabolic and physiological functions [1-4]. Haloperoxidases are probably involved in the formation of many of these organohalogens.

Like other peroxidases, haloperoxidases may contain haem as a prosthetic group, e.g. chloroperoxidase from *Caldariomyces fumago* [5]. However, a large group of iodoperoxidases and bromoperoxidases have been discovered containing 1 mol vanadium/mol enzyme, which was shown to be essential for peroxidase activity. These vanadium-containing haloperoxidases were isolated mainly from seaweeds [6–8], although a vanadium bromoperoxidase was also discovered in the terrestrial lichen *Xantheria parietina* [9]. EPR studies [10] and K-edge X-ray studies [11, 12] on vanadium bromoperoxidase from the brown seaweed *Ascophyllum nodosum* showed that the oxidation state

Correspondence to R. Wever, E. C. Slater Institute, University of Amsterdam, Plantage Muidergracht 12, NL-1018 TV Amsterdam, The Netherlands

Fax: +31 20 5255124.

Abbreviations. RT, reverse transcription; CPO, chloroperoxidase; BPO, bromoperoxidase.

Note. The novel nucleotide sequence data published here have been submitted to the EMBL sequence data bank and are available under accession number X58369.

of the metal was vanadium(V). Apparently, the redox state of the vanadate does not change during catalytic turnover. A model has been proposed in which the vanadium(V) coordinates the hydrogen peroxide. The resulting enzyme-peroxo intermediate is able to oxidize bromide to hypobromous acid [12, 13].

Recently, it was shown that certain terrestrial fungi also produce and excrete vanadium chloroperoxidases [14]. The extracellular vanadium chloroperoxidases produced by these fungi may be partly responsible for the formation of naturally occurring organohalogens, which have been found in the terrestrial environment [15]. It has been suggested that the fungal vanadium chloroperoxidases function in degradation and/or penetration of plant tissues (unpublished results), but the precise function of these haloperoxidases has yet to be established.

Two vanadium chloroperoxidases from the fungi *C. inaequalis* and *Drechslera biseptata*, have been biochemically characterized in our laboratory in detail [14, 16, 17]. These enzymes were found to share some interesting features. Besides the ability to form free hypochlorous acid, they display remarkable thermostability and resistance towards denaturation by organic solvents and detergents [17]. Furthermore, in contrast to haem chloroperoxidase from *C. funago* [5], they show a considerable tolerance to the hypohalous acid product [18]. Finally, when grown in liquid cultures in the absence of vanadate, the fungal vanadium haloperoxidases are secreted into the medium as an apoprotein, which can be easily isolated. This inactive apoform can be reconstituted to full activity when vanadate is added [14].

Although vanadium haloperoxidases constitute a large group of enzymes, little is known about their structure and function. In this study, we describe the cDNA sequence determination of a vanadium chloroperoxidase from the fungus *C. inaequalis* in order to elucidate the primary structure and to initiate a preliminary study of the gene expression. The sequence is compared with other protein sequences to study if any structural relationship exists.

EXPERIMENTAL PROCEDURES

Organisms. The fungi C. inaequalis and D. biseptata were obtained from the Central Bureau voor Schimmelcultures (CBS strain nos 102.42 and 371.72, respectively; CBS, Baarn, The Netherlands). Escherichia coli strain HB101 was used as a host for pUC18/19 derivatives.

Isolation of the chloroperoxidase and protein sequencing. The chloroperoxidase was isolated and purified from liquid cultures of *C. inaequalis* as described previously [14], except that after DEAE chromatography two additional purification steps were performed using an FPLC system (Pharmačia LKB). First a phenyl-Sepharose CL-4B hydrophobic interaction column was used to bind the enzyme in the presence of 2 M NaCl in 50 mM Tris/HCl, pH 8.3, followed by elution with one column volume of 0.1 M NaCl in 50 mM Tris/HCl, pH 8.3. For the final purification, a Mono-Q HR 5/5 anion-exchange column was used to bind the enzyme, followed by elution with a gradient of 0− 0.5 M NaCl in 20 mM piperazine/HCl, pH 5.4. Subsequent concentration of the enzyme was carried out using rotation evaporation, followed by dialysis against 50 mM Tris/SO₄, pH 8.

The purified chloroperoxidase was enzymically digested with the proteases *Staphylococcus aureus* V8 and trypsin or chemically cleaved with CNBr [19]. The resulting peptides were separated using SDS/PAGE according to Laemmli [20] or on a Tricine gel according to Schägger and Von Jagow [21] and subsequently blot-transferred to poly(vinylidene difluoride) membranes (Immobilon-P, Millipore) using Caps transfer buffer [10 mM 3-(cyclohexylamino)-1-propanesulfonic acid, 10% methanol, pH 11] as described by Matsudaira [22]. After electrophoretic elution, the membrane was rinsed for 5 min with deionized water, stained with 0.1% Coomassie brilliant blue R-250 in 50% methanol for 5 min, and destained in 50% methanol, 10% acetic acid, for 10 min at room temperature. The membrane was finally rinsed in deionized water for 10 min and air dried. Peptide bands were submitted to automatic Edman sequencing on a Porton LF 3000 protein sequencer (Beckman Instruments).

Enzyme activity assays. Chloroperoxidase activity was assayed in filtered culture media. Samples were concentrated using Centricon-30 concentrators (Amicon). The chloroperoxidase activity was assayed by measuring the chlorination of 50 μ M 1,1-dimethyl-4-chloro-3,5-cyclohexanedione ($\epsilon = 20.2 \text{ mM} \cdot \text{cm}^{-3}$ at 290 nm) [5] using a Varian Cary-17 spectrophotometer. The activity assays were carried out in 0.1 M sodium acetate, pH 5.5, 5 mM potassium chloride, and 1 mM hydrogen peroxide. Qualitative enzyme activity assays were performed using a phenol red assay as described by De Boer et al. [23]. Chloroperoxidase activity on SDS/PAGE gels [20] was detected by soaking the gel in 1 mM *o*-dianisidine, 100 mM potassium bromide, and 100 mM potassium phosphate, pH 6.5, and H₂O₂ to a concentration of 1 mM. The SDS present during electrophoresis does not interfere with chloroperoxidase activity.

Analyses for metal content. Metal analyses of vanadium chloroperoxidase samples were performed on a Hitachi 180-80 Zeeman polarized atomic absorption spectrophotometer using a Hitachi pyrolysis graphite cuvette.

Determination of protein disulphide bridges. Protein disulphide bridges and free thiols were determined using a method described by Thannhauser et al. [24]. For further identification of putative disulphide bridges in the chloroperoxidase, SDS/ PAGE using a 10% gel according to Laemmli [20] was performed. The chloroperoxidase sample mixtures were boiled with or without 2-mercaptoethanol (5%) for 5 min before applying the samples to the gel. After gel electrophoresis, the gel was stained with 0.1% Coomassie brilliant blue R-250 in 25% methanol, 10% acetic acid, for 1 h and destained in 25% methanol, 10% acetic acid. Differences in the running pattern indicated the presence of disulphide bridges within the protein.

Assay for glycosylation of the enzyme. To determine the presence of N-linked sugars in the protein, the purified vanadium chloroperoxidase was treated with N-glycanase enzyme (Sanbio). Changes in mobility of the N-glycanase-treated enzyme when compared to the untreated enzyme were tested using SDS/PAGE according to Laemmli [20]. The presence of glycan residues was further tested using a sensitive antibody-based dioxygen glycan detection kit (Boehringer Mannheim).

Isolation of mRNA, cDNA synthesis, cloning and analysis. For the isolation of RNA, spores of C. inaequalis were inoculated directly into a fermentation medium containing 4 g veast extract and 2 ml microelement solution/l [14]. After several days of growth, the mycelia were harvested by filtration and lyophilized. The lyophilized C. inaequalis mycelia were ground under liquid nitrogen. RNA was extracted by adding an RNA extraction buffer (42 mM sodium citrate, pH 7, 0.83% N-laurylsarcosine, 50 mM 2-mercaptoethanol, 1% Triton X-100 and 4 M guanidine isothiocyanate), and incubating for 1 h at room temperature. 0.1 vol. 2 M sodium acetate, pH 4, and 1 vol. phenol/chloroform/isoamvl alcohol (25:24:1, by vol.) was added and the mixture placed on ice for 15 min. After centrifugation for 10 min at $10000 \times g$, 4°C, the aqueous phase was collected, 1 vol. absolute ethanol was added and the mixture was incubated for 1 h at -20° C followed by brief centrifugation at $10000 \times g$. The pellet was resuspended in an appropriate volume of RNA extraction buffer and fractionated by ultracentrifugation in a gradient of cesium chloride [25]. The RNA pellet was carefully washed

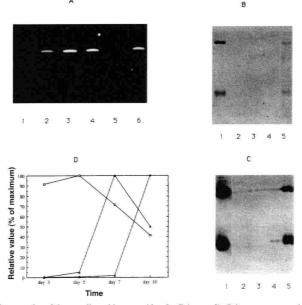


Fig. 1. Growth-stage-specific expression of the vanadium chloroperoxidase by *C. inaequalis*. Cultures were grown for the indicated time periods. (A) RT-PCR result using 5 µg isolated total RNA. Lanes 1-4, days 3, 5, 7, and 10, respectively; lane 5, empty; lane 6, control PCR reaction using the same primers on a cloned cDNA gene fragment. (B) SDS/PAGE gel stained with Coomasie brilliant blue. Lane 1, control (purified vanadium chloroperoxidase); lanes 2-5, concentrated culture medium of days 3, 5, 7, and 10, respectively. (C) Similar SDS/PAGE gel as shown in B after staining for chloroperoxidase activity with o-dianisidine in the presence of bromide (see also Experimental Procedures section). (D) Total dry mass of the fungus in the cultures (\Box), enzyme activity (1,1-dimethyl-4-chloro-3,5-cyclohexanedione chlorination) (\Diamond) and competitive RT-PCR results (Δ) (using 5 µg isolated RNA per culture) on days 3, 5, 7, and 10. The results are in each case relative to the highest value observed.

and stored in a 75% ethanol solution at -70° C. For mRNA isolation the RNA was precipitated and resuspended in ribonuclease-free water after which the mRNA was extracted using the poly(A) tract mRNA isolation kit (Promega).

Four 20-residue degenerate oligonucleotides were designed based on the amino acid sequences of chloroperoxidase peptides and used as primers in polymerase chain reactions with firststrand cDNA from C. inaequalis as a template. First-strand cDNA synthesis, used in polymerase chain reactions, was carried out on isolated mRNA from C. inaequalis using the Pharmacia first-strand cDNA synthesis kit (Pharmacia). Polymerase chain reactions were performed using a thermocycler (Eppendorf mastercycler 5330) and Taq polymerase (Promega). For optimal amplification of the chloroperoxidase-encoding cDNA using the degenerate primers, the polymerase chain reaction (RT-PCR) was performed at 46°C for 30 cycles. The two resulting specific fragments, 71-bp and 62-bp long, were ligated into a pUC18 vector, cloned and sequenced from both strands using the Taq Dyedeoxy Terminator Cycle sequencing kit (Applied Biosystems) on a model 373A sequencer (Applied Biosystems). Based on the DNA sequence results, new primers were designed and used for polymerase chain reactions in RT-PCR to obtain the complete encoded region (see also the Results section), or as primers in DNA sequencing for complete double-stranded sequencing of the resulting clones.

To obtain the 5' region of the mRNA encoding the chloroperoxidase the 5'-Amplifinder RACE kit (Clontech) was used. The gene-specific primers used in this method were based on DNA sequence data obtained as described above.

Quantification of chloroperoxidase mRNA levels. To obtain semi-quantitative information of chloroperoxidase-encoding mRNA levels in the fungal cells at different stages of growth, competitive RT-PCR [26] was used. Total RNA was isolated from the cells of interest, 5 μ g of which was used for reverse transcription as described above. Competitive RT-PCR was carried out using 1 μ l of a 1:100 dilution of the first-strand DNA mixture as a template and two gene-specific oligonucleotides as primers for the amplification reaction. An internal control PCR fragment of known size and concentration, containing the same primer template sequences, was added to the PCR mixtures in different concentrations. The amount of this competitor DNA yielding an approximate equimolar amount of product as the target cDNA was taken as a measure for the amount of target mRNA.

Isolation of genomic DNA and analysis. C. inaequalis genomic DNA was isolated from lyophilized mycelia ground under liquid nitrogen and extracted with an appropriate amount of extraction buffer (200 mM Tris/HCl, pH 8.5, 25 mM EDTA, 250 mM NaCl, 1% SDS and 0.2 mg · ml⁻¹ proteinase K). After incubation overnight at room temperature, 0.7 vol. phenol and 0.3 vol. chloroform were added and mixed vigorously. The tubes were centrifuged at $10000 \times g$ and the aqueous layer was transferred to a clean tube. The genomic DNA was precipitated with 2 vol. absolute ethanol. After centrifugation for 5 min at $5000 \times g$

Table 1. Amino acid sequence data. N-terminal amino acid sequence of five peptides from the chloroperoxidase of *C. inaequalis.* The peptide fragments were obtained by cleavage with trypsin, V8 protease, or CNBr and subsequently separated with PAGE, blotted onto poly(vinylidene diffuoride) membranes and analysed using an amino acid sequencer. Sequences used to design degenerated primers for amplification of DNA (see text) are underlined. The numbers of the first and last amino acid residue of each peptide indicate the position in the derived protein sequence (see also Fig. 2).

No.	Amino acid sequence	Cleaved/digested with
1	12 IPDEPEEYN 19	trypsin
2	109 MLSSLYMKPVEOPNPNPGANISDNAYAOLGLVLDRSVLEA 149	CNBr
3	183 GYHPTPGRYKFDDEP 198	V8 protease
4	251 SNADETAEYDDAVRVAIAMGGAOALNSTK 280	trypsin
5	450 DLRQ(S)YDPTAPIEDQPGIVRT 471	trypsin

the pellet was resuspended in 2 ml 10 mM Tris/HCl, pH 8.0, 1 mM EDTA buffer and treated with RNase (Boehringer Mannheim) as recommended by the manufacturer. The genomic DNA-containing solution was extracted with phenol/chloroform/ isoamyl alcohol (25:25:1, by vol.) and, after ethanol precipitation, finally dissolved in a suitable volume of 10 mM Tris/HCl, pH 8, 1 mM EDTA.

For Southern-blot analysis of the genomic DNA, the DNA was digested with several combinations of restriction enzymes and after agarose gel electrophoresis blotted onto a nitrocellulose membrane [25]. Hybridization was carried out using a radiolabelled gene-specific fragment, which was made by primer extension using primers 1 and 4 (Fig. 2) and first-strand cDNA as a template. The resulting fragment was labelled by random priming using a^{-32} P-labelled dATP [25]. Based on this result, a minilibrary was made using genomic DNA digested with *Psrl*, which was inserted in the vector pUC18. The library was screened with the same probe as described for the Southern blot. A positive clone was isolated and sequenced from both strands to confirm the cDNA sequence results.

RESULTS

Regulation of chloroperoxidase gene expression. As already shown by Van Schijndel et al. [14], the chloroperoxidase enzyme could only be detected in liquid cultures after several days of growth. Furthermore, we found that the production of the enzyme was retarded for several days when rich growth media were used. To investigate this finding, the expression of the chloroperoxidase-encoding gene in relation to the growth stage of the fungus was studied. Cells were grown in separate cultures and the chloroperoxidase activity in the medium and the chloroperoxidase mRNA content in the fungal cells were monitored after 3, 5, 7, and 10 days of growth. The results are depicted in Fig. 1. At day 3, no chloroperoxidase mRNA could be detected, while a maximum chloroperoxidase-encoding mRNA yield was obtained after 7 days of growth. The increase in chloroperoxidase-encoding mRNA coincides with a decrease in total dry mass of the fungal culture. Chloroperoxidase activity was not detectable after 3 days and was still very low at day 7. At day 10, however, the chloroperoxidase activity increased 60 times compared to day 7. The increase in activity correlates with the amount of the vanadium chloroperoxidase as observed by SDS/ PAGE according to Laemmli [20].

Determination of peptide sequences for *C. inaequalis* chloroperoxidase and designing DNA probes. FPLC purification of the chloroperoxidase protein showed a single peak after the final purification step using a Mono-Q anion-exchange column. Isoenzymes for the chloroperoxidase were not found. Sequencing of the purified chloroperoxidase protein revealed that the protein was N-terminally blocked. After cleavage of the protein with trypsin, V8 protease, or CNBr, five peptide fragments were partially sequenced. The results of the amino acid sequence determination are summarized in Table 1.

Nucleotide sequence of the cDNA. Based on the amino acid sequences of the peptides, four degenerate oligonucleotides were designed (see also Table 1). Using first-strand cDNA as a template and the degenerate primers, the regions encoding peptides nos 2 and 4 (Table 1) were amplified. The small amplification products were cloned into a pUC18 vector and the sequence was determined. Based on these sequence results, two specific primers, 1 and 4 (Fig. 2), were designed and used to amplify the intervening region with RT-PCR. The specific fragment was cloned into pUC18 and sequenced. Two primers, 4 and 5 (Fig. 2), based on the 5' region of this 500-bp fragment, were used as gene-specific primers in the 5'-Amplifinder RACE protocol to obtain the 5' encoding region of the mRNA. After RT-PCR the resulting fragments of about 750 bp were cloned into pUC18 and sequenced. To obtain the 3' coding region of the mRNA, primer 3 (Fig. 2) and the T₁₈ Not primer for first-strand cDNA synthesis were used in RT-PCR. The resulting fragment of about 1400 bp was cloned into the pUC18 vector and subsequently sequenced. The three overlapping cDNA clones were all sequenced at least three times from both strands using the cloned PCR fragments derived from three independent amplification reactions in order to exclude possible sequence mistakes due to the incorporation of incorrect bases during the amplification reaction. If necessary, internal sequencing primers were designed and used to obtain a complete double-stranded sequence.

The complete cDNA sequence and its derived amino acid sequence is presented in Fig. 2. The protein sequence is depicted starting with the first methionine residue. The cDNA sequence comprises 2091 bp plus a poly(A) tail and contains an open reading frame that encodes for 609 amino acids. The sequence AATACA, 58 bp upstream of the poly(A) tail, is very similar to the eukaryotic polyadenvlation signal AATAAA [27].

To investigate if there are multiple genes encoding for isoenzymes of the chloroperoxidase, a Southern-blot experiment was carried out with differently digested genomic DNA (Fig. 3). The blot was hybridized using a radiolabelled 500-bp PCR fragment, obtained with primers 1 and 4 (Fig. 2) and cDNA as template, as a probe. The result (Fig. 3) shows that only one specific DNA fragment reacts with the chloroperoxidase gene-specific probe, confirming the absence of isoenzymes. Using the results obtained with this Southern blot. a *PstI* fragment, containing the encoding DNA sequence was determined on both strands to confirm the cDNA sequence results (Fig. 2).

Translation of the full open reading frame gives a protein of 609 amino acid residues with a calculated molecular mass of

gttggc<u>tataaa</u>gcctgtgagct

ttetegttetatateagaateacaatteacttgeteagtetacteactactactacaateacacacattateacatteacg

6. AAC TAC ATA CTA TTC TOT DAAC CAT TTC OUT TT O GA CTC AAC DOC OT A AT CAC AT 21 Am Try 11e fee Phe Try Am His Val Ciy Leu Cis Leu Am Art VAL Thr His The 21 OR DAG CC CTC AC DOL DAG CAC CT CTC TTC OC CALD DAT CTC OUT OF CAT CTC 31 Det His Phe Try Am His Val Ciy Leu Cis Leu Am Art VAL Thr His The 31 Con Cir CC CTC AC DOL DAG CAC CT CTC TTC OC CALD DAT CTC OUT OF CAT CTC 41 Det His Pha 1a Try The Fire Try CT ATC TCC CTC CCG ACC ATC TTC ACA DC TC CTC CTC CTC 31 Let His Pha 1a Try The Fire Tity CT ATC TCC CTC CCG ACC ATC TTC ACA DC TC CTC CTC CTC 31 APT THT OUT DAG CTC CTC TTC CTC CTC CTC ACC ATC CTC ACC ATC CTC C										~						200		and		107	100
21 Am Tyy 11e Lee Phe Typ Am His Val Cly Leu Clu Lee Am Art Val Thr His The Ise The GA Core Cro AGO GA CCA CT TC CT CT CT CAC GAT GAT CT CAC TT GA CAC CT CT CAC TT GA CAC CT CT CAC TT GA CAC CT CAC TT C	1	Mat	Gly	Ser	Val	Thr	Pro	Ile	Pro	Leu	Pro	Lys	Ile	Asp	Glu	Pro	Glu	Glu	Tyr	Asn	Thr
 41 diy diy Pro Leo Thư điy Pro Pro Leo Ser Alia Arg Alia Leo Tiy Met Leo Hi Met Leo Hi S diy Giy Pro Leo Thư Chi Try Cri Try Try Try Try Cri Cri Cro Acc And Cri Cri Cri Cri Cri Cri Cri Cri Cri Cri		AAC	TAC	ATA	CTA Leu	Phe	TGG Trp	ARD	CAT	orc Val	GGT	TTG Leu	GAA	Leu	AAC Asn	CGC Arg	GTA Val	ACT Thr	CAC	ACT	GTT Val
61 114 HIS AND ALS TYPE HRE BET 116 CYPE HTO HEAD HTM AND PHE HTM THE PHE LEAD BET 321 AND ACT AND AND COLORD THE COT CYPE AND CYPE AND COT CAD AND AND AND TOT CYDE 31 AND THE CUL AND ALS ALS TYPE AND AND CYPE TYPE AND CUL AND CYCLE THE CUL AND AND CYPE THE 31 AND THE CUL AND ALS ALS TYPE AND AND CYPE TYPE AND CUL AND CYCLE THE CUL AND CYCLE THE CUL AND ALS CYLE AND CYCLE AND CYPE AND FYPE AND PYPE AND		GGA Gly	GGC Gly	CCC Pro	CTG Leu	ACG Thr	GGA Gly	CCA Pro	CCT Pro	erc Leu	TCT Ser	GCC	AGG Arg	GCT	CTG Leu	GGT GIY	ATG Met	CTG Leu	CAC His	TTG	GCT Ala
 Any Thr Glu Am Ala Ala Tyr Arg Leu Pro Ser Pre Am Gly Ala Am Arp Ala Arg Ala Cai Gro Ger Gal Gut Ger Cre Ala And Tro Tro Tor Tro. Hor Tar Lar And Cai Ger Gel Gal Gut Ger Cre Ala Arg Cre Ala Cort Ala Cort Gal Gut Ger Cre Ala Arg Tor Unit Tar Tar Cai Cr Tar Ger Cai Cai Ger Gal Gal Gut Ger Gel Cai Ala Arg Tor Cort Gal Cr Tar Ger Tal Gut Gal Gal Gal Gal Gal Gal Gal Gal Gal Gal		ATT Ile	CAC His	GAC Asp	GCC Ala	TAC Tyr	TTT Phe	TCT Ser	ATC Ile	TGC Cys	CCT Pro	CCG Pro	ACC Thr	GAC Asp	TTC Phe	ACC Thr	ACC Thr	TTC Phe	CTC Leu	TCA Ser	CCT Pro
 101 Ala Val Ala Gly Ala Ala Deu Dy Met Leu Ser Ser Leu Tyr Met Luy Fro Val Glu 101 CT ANG COT ANG COT GOLG COT ANG ANG COT ANG COT TAT COT TA GLY TAT GLY TAN GLY TEA AND FROM AN FOR GLY ALA AND TALE SER AND AND AND AND AND AND AND AND AND AND		GAT Asp	ACT Thr	GAG Glu	AAT Asn	GCC Ala	GCG Ala	TAC Tyr	CGT Arg	CTA Leu	CCT Pro	AGC Ser	CCT Pro	AAT Asn	GGT Sly	GCA Ala	AAT Asti	GAT ASP	GCT Ala	COC Arg	CAA Gln
1121 Pro Am Pere Am Pere Clip Ala Am 11e Ser Amp Ami Ala Tyr Na Gun Dec Clip Jee Clip Ame Can Pere Clip Ala Ami Ala Tyr Tak Gun Dec Clip Jee Clip Ame Ami Ala Tyr Tak Gun Dec Clip Ame Ami Ala Tyr Tak Gun Dec Clip Ame Ami Ala Tyr Ala Gun Dec Clip Ame Ami Ala Tyr Ala Gun Dec Ami Ame Ami Ala Ser Tam Ani Clip Ame Ami Ala Tyr Ala Gun Dec Ami		GCA Ala	GTC Val	GCT Ala	GGA Gly	GCT Ala	GCC Ala	CTC Leu	AAG Lys	ATG Met	CTG Leu	TCT Ser	TCA Ser	erc Leu	TAC Tyr	ATG Met	AAG Lys	ccc Pro	GTC Val	GAG Qlu	CAG Gln
341 Lee, Aap Arg Ser Val Lee Cli Alta Pro Cly Cly Val Aap Arg Cli Ser Ala Ser The 441 The Out Dad Dat Own doc Own or the The Oack of the Che Cla Cl Cut of Out 55 5 Pho Cly Clu Aap Val Lea Aap Val The The Ala Aat The Cle Che Che Che Che Che Che Che Che Che Ch								GCC Ala	AAC Asn	ATC Ile	TCC Ser	GAC Asp	AAC Astr	GCT Ala	TAT Tyr	OCT Ala					GTT Val
161 Phe Gly Glu App Val Ala App Val Phe The Ala Eau Leu Ann App Pro Arg Gly Ala S41 CAN GAO GOT TAC LAC COT ACC ACC GGC GGC TOT AAA ATT GAC GAT GAA COT ACT CAN S41 CAN GAO GGC TAC CAN COT ACC ACC GGC GGC GGT TAT AAA TTT GAC GAT GAA COT ACT CAN S41 GLU GLU TAY THE HE TOT PRO VAL AND TAY TAY TAY THE AND AND COT THE GG S41 VAL Leu LE PTO VAL AND FOR AND AGT TOT GGT GGC AAA AND COT THE COT S41 VAL Leu LE PTO VAL AND FOR AND AGT COT ANT GGT COC ANA GGC COT THE CAN S41 VAL Leu LE PTO VAL AND FOR AND AGT COT ANT GGT COC ANA GGC COT THE COT S41 VAL Leu LE PTO VAL AND FOR AND AGT COT AND GGT COC ANA GGC COT THE COT S41 VAL Leu LE PTO VAL AND FOR AND AGT COT AND GGT COC ANA GGC COT THE COT S41 VAL Leu LE PTO VAL AND FOR AND AGT COT AND GGT COC ANA COC ANA CANC AND COT THE COT S41 VAL LEU LEA AND FOR TOT G1 LEU ANT GGT GAC GAC ACC COC GAG TAT GAC GAC S41 LEU ALA AND FOR TOT G1 LEU ANT GGT GGT COC ANC COC ACC AND GOT GAC S41 LEU ALA AND FOR TOT G1 LEU ANT GGT GGT COC ANC COC ACC AND GGT AT S41 AND COC GAC CTAC GCC ATU GGT GTT COC GCT CAC GGT GGT COC ACC CAC GGT GGT ACT S41 TOO COA ACA COA COA GGC COT TOT ACT GGC COC ACC AND GGT CAC ANC COC ACT GGT COA S41 TOO COA ACA COA COA GGT COT TAC COT COC ACT GGT GGT COC ANC COC COT GT GAC S41 TOO COA ACA COA COA GGT GGT TAC COT GCC ACT GGT GGC COC COT GT GAC S41 TOO COA ACA COA GGA GGT ACC ACC ACC GGT GGT GGC COC COT GT GAC S41 TOO COA ACA ACC ACC GAC GGT GGT GGT COC ACT COT COC COC COT GG S42 ALA COM COA ACC ACC GGT GGT TOT GAT GGT COC COT COT COC COC COT GG S44 ALA COM COA ACC COT GGT ACC AGA GGA GGT COA COC COT COT COC COT GGT GGC S44 ALA COM COA GAC GGC GGT GGT GGT GGT GGT GGT GGT GGT GG		CTC Leu	GAC Asp	CGA Arg	TCA Ser	STT Val	CTG Leu	GAG Glu	GCA Ala	CCT Pro	GOT Gly	GGC Qly	GTG Val	GAC Asp	CGA Arg	GAG Glu	TCA Ser	GCC Ala	ACT Ser	TTC Phe	ATG Net
 SHI CAD GAD GROT PAC CAD COT ACK GOD GROT PAC AAA TITU GAD CAD GAD ACT ACT GAD CAD SHI GUL GUP TYP HILE PTO VAL AND TTO PTO FYP LEVE PTO AND AND PAD GUL PTO THE HILE GID GTU GTU TYP TYP THE PTO AND AND PTO AND AND AND GUL PTO THE HILE GID GTU GTU CAT ATT CAD GAD CAD CAD AND PTO FAN GUL PTO THE LEVE PTO VAL AND PTO AND AND PTO AND GUT CTU CAD AND CAD CAD CAD CAD CAD CAD CAD CAD CAD CA																					
ALL GUT OF CHE ATT COA GTA GAR CHE AND CHE		CAG	GAG	GGC	TAC	CAC	CCT	ACA	ccc	GOC	CGC	TAT	AAA	TTT	GAC	GAT	GAA	CCT	ACT	CAC	CCT
301 Val Val Lee TLE PTO Val Asp PEO Ami Am PEO Ami CUI PEO LYM MEE PEO PER Arg 461 TAC CAC DEC CEA TRO TAC DEC ADA ACC ACO AAG DET TET DET ACT PAGE 461 TAC CAC DEC CEA TRO TAC DEC ADA CAC ACO AAG DET TET DET ACT PAGE 461 TAC CAC DEC CEA TRO TAC DED ADA CEA ACO AAG DET TET DET ACT PAGE 461 TAC CAC DEC CEA CEC DED CEN CET TET NAT GED ANC AGA ACO CEG AGA TAT DAG AGA 461 TAL ALLA ALLA PEO DE PEY DEL ARG SER AMA ALLA AGA CEC CAC AGA DET AGA 461 TAL ALLA ALLA PEO DE ALLA DES ANT GED TAC DAG AGA CAC CEG AGA CEA AGA DET 461 TED CAC DAG CEA CEC DE DEC CEN DET TET AAT GED ALC GAA ACO CEG CAC AGA DET AGA 461 TAL ALLA AME TEL VALUE ARG SER AMA ALLA AGA CEC CE ACC CAC AGA DET AGA 461 TED CAC DAG CEA CEG DEC CEN DET TET ALLA SEN AME DET THE LYM ARG SER 461 TED CAC DAG CEA CEG DEC CEN DET TET ALLA SEN AME DEL THE LYM ARG SER 461 TED CAC DAG CEA CEG DEC CEN DE TAT TAL DE ALLA SEN AME THE LYM ARG SER 462 TET DEC THE ALLA DAG CAL DEC THA TAC DE DEC TAC DET GED ATC CEA CEC CEC CET AGA DE ALLA SEN DEA CEA CAC AGA DE CET ATC DE CEA DE CEA DE TAC DETA AME ALLA DAMA DEM 462 TET DEC THE ALLA DAG CAL ATC DETA COT DEC ATC DEA CEA DET CEA DE CEA CE CEA DE TAC 463 TET DEC THE ALLA DAG CEA ATC DETA COT DEC ATC DEA CEA DE TAC DE ATC AMA AMA DAMA ADA 464 TET DE CEA DAG AGA CET OFT TET DE SE DE LA VAL TET DETA DAG AGA DETT 464 DE ALLA SEN SER DE UL VAL TATA AMA ATC DE DAG THE DEC CEA DE TAC DATA DAG AMA ATTO 465 DE TET ALLA DAG DETA TET DETO TED CAA DE CAA ATTO DE ATC DAG ATTO 464 DE ALLA DAG DE CEA DE DAG THE DE ALLA ATTO DE ATC DAG ATC TET DE DAG DE 465 DE CEC CEA DAG AGA CET OFT ACC ATC AGA CAA CEA DE TET DAG DE TET DAG 466 DE TET DEG DET CEA DEC ACT AGA ACA CEA CEA CEA TET DE ATC DAG DET TEA. TET TET DE 466 DE CEC DET CEA DE DEC AGA DEC CET TET DAG DETT TET CEA ATTO TET DE 466 DE CEC DET CEA DE DE CEA DEC CEA DE ATT DE DE TER DE DE TER DE DE TER DE DE TER DE																					
321 Tyr His Ala Fro Pie Tyr Cily Lys Thr Thr Lys Arp Phe Ala Thr Cin Ser Cili Ala Tyr Ho Got GAC CAL Col Gib Criti Cer Cer An Got GAC GAC AC Col Gib Cat Cat Ga Ser Col Gat Cat Cer Act Col Gib Criti Cer Cer An Got GAC GAC AC Col Gib Cat Cat Ser Col Grit Cer Arc Col Arb Gat Gar Gar Cat Cat Cer Cer Act Cer Act Cer Act Cer Arg Ser Yei Arg Yei Ala His Lie Ala Her Child Yily Ala Gin Ala Lee And Ser Thr Lys Ang Ser Yei Arg Yei Ala His Lie Ala Her Child Yily Ala Gin Ala Lee And Ser Arc Cer Art Got Act Arg Ser Tyr Ber Thr Act Cer Arb Get Gar Gar Cer Ce Act Cer Act Cer Art Got Act Ser Cer Cer Thr Cer Arb Cer Arb Get Gar Gar Cer Cer Cer Cer Cer Cer Cer Art Got Act Ser Cer Cer Thr Cer And Cer Arb Get Gar Gar Cer Cer Cer Cer Cer Cer Cer Cer Art Ser Ser Yei Ala His Lie Ala Her Ala Yei Ala Styr Har Gily Ser Ann Lee Lie Gily Thr Ser Cer Thr Act Act Cer		Val	Val	Leu	Ile	Pro	Val	Asp	Pro	ABI	Asn	Pro	Asn	Gly	Pro	Lys	Met	Pro	Phe	Arg	Gln
 Les Ala Anp Pro Pro Dly Les Arg Ser Asn Ale Anp Glu Thr Ala Glu Tyr Asp Anp 751 Grc GG GTG GT ATC GG ATG GT ATG GG ATG GT GT GT CA AGT GT AAG TG AGC AGG AGG AGG AGG AGG AGG AGG AGG AG	221	Tyr	His	Ala	Pro	Phe	Tyr	Gly	Lys	7hr	Thr	Lys	Arg	Phe	Ala	Thr	Gln	Ser	Glu	His	Phe
361 yell Arp Yell All The AllA Net Diy Diy Alla Gin Alla Leu Ann ser Thr Lye Arp Ser Art Dig CAG Arb GCA CAG GGE TTA TAC TOG GCC TAC GAT GGE TCA ARC TTE ATT GGE ACA SER SER AND LEU DIY THE ALLA SIN DIY AND UNE AND TACA YAR ARD ARC ARC CAG GGE CTA TAC TOG CAC TAC GAT GGE ATA CAC TTA GGE ACA SER SER AND LEU DIY THE ALLA SIN DIY AND UNE AND ARC TAC GT CAC TAC GAT GGE ATA CAC TAC GAT GGE ATA CAC ARG GAT ATT GGE CAC TTA GGE ATA CAC ARG GAT ATA CAC ARG GAT ATT GGE CAC TTA GGE ATA CAC ARG GAT ATA CAC ARG GAT ATT GGE CAC TTA GGE ATA CAC ARG GAT ATA CAC ARG GAT ATT GGE CAC TTA GGE ATA CAC ARG GAT ATA CAC ARG GAT TTAC CAC CAC CAC ARG																					
781 THE GLM THE ALS GLM GID LEW TYP THE ALS TYP ANP GLY SEE AND LEW ILE GLY THE 901 CTO COT THE TAKE AND CAS AND CAY COT COT CAT CAS GTM CAS TAK CAS AND AND AND CAS AND CAS AND CAS CAT CAT CAT CAS GTM CTA CAS AND AND CAS AND CAS THE VAL ANY LINE ILE ALL VAL THE TYP LINE GLAY GLM GLA GLAY CAS AND CA		Vai	CGC Arg	GTC Val	GCT Ala	ATC Lie	GCC Ala	ATG Met	GGT Gly	GQT Gly	GCT Ala	CAG Gin	GCT Ala	CTC Leu	AAC Asn	TCC Ser	ACC Thr	AAG Lys	CGT Arg	AGC Ser	CCA Pro
301 PPC Arg PPC Tgr Am GLE 118 Val Arg Arg 116 Ala Val Thr Tyr Dor Dyr GLE GLU S41 GTT GCC AAC ACC GAA GTC AAC AAT GCC GAT TTO GCC CGC TCT GTC GCC GCC GTC GTC GCC GAC GAA GTC GAA GTC AAC AAT GCC GAT TTO GCC CGC TCT GTC GCC GCC GTC GTC GLU Lau Ala RES Ser GLU VAL Alam Ann Ala Ala PPC Ala Arg He Cha Arg TCG GCC TTO GCC ALC GAT GTC ALAM STAR Ser GLU VAL ALAM ANA GCC GAT TTO GCC GAC TTO TTO GCC GTC GTC GTC GLU Lau Alam Ser GLU VAL GTC GTC ACT GTT TTO GTC GA AAC ATT GGL AT TTO GAD TTO GLU HE GLU HE TTO 1361 Ala Cys Thr Asp Ala GTY GTC ANG GAC GTC GCA AAC AAT GGT GLU TTO TGO GTC 361 PTC Lau B8C GIY VAL ATS AAR ANA JUL ATG FCD AAR HIS GLY AND FTC GAT FCT GAT GTC 361 PTC Lau B8C GIY VAL ATS AAR AND JUL ATG FCD AAR HIS GLY AND FTC FTT TTO GC GTC 361 PTC GGT GCC GCA GCT ATT AAC ACC AAC GAC ATT CCA HCT AGG CT 011 PTC Set GCC CCA GCT ATT AAA ACA CAC GAC ATT CCA HTC AAG GTC GCT GTC ATA TA 401 GIY ATS YAL BY ALA FTC ALA HAA ATA ANA JUL ANA THE GLU ANG VAL GTT GAT ATT ATA 411 GIY ATG GAA GCT GAA GCG GAT GTT GGC GAC GTT CCA HCT AGG CT 412 GIY ATG YAL GUT HTT TTE LAY ANG ANG GAC GAC GCA GAC ACT TC GCA GCT TTO GAC ATG ATA 413 GIY ATG GAA GAC GAA GCC GGT AACC GAC GAC GAC GAC GTT GAC TTO GCC ACT TC GAA GCT 414 GLU ALB GTC GAA GCC GGT GAA GCC GGC GAC AAC ATG TTO GCC TAT GAT GCC 415 SET GLU GLU ALB GTT GAA GCG GGT GAC ACT GGC GAC TAT GAC GTT GAA TAT GAC 416 GAA GAC GAA CCC GGT AACC GGT GAC CCC GCC ATT GTT GCC GAC TTO GAC GCC GAC 417 SET GLU GLU ALB GLU ATA HIS ILLE SET ATS JLE HILL ATS HILF HAA FL AND SE FALL 1414 GAA GAC GAA CCCA GGT AACC GGC GAC ATT GAT GCC GAC TTO GAC TOG GCT 418 SET GLU GLU ALB GLU ATA HIS ILLE SET ATS JLE FIN HILE ANA GAU GCT TTO TT 418 GLU ALB GAA CCA ACC GAC TTO GCC GAC ATT GAC GCC GTC TTC 411 GLU LAU AGE GAA GCC GAC ATT GGC GAC TTC GCC GAC TTO GCC GAC TTC 411 GLU LAU AGE GAC GAC GAC GT ATT GGC GAC TTC GCC GAC TTC GAC GGC GTC TTC 511 ALB ANA ALB ALB ATG ANG FLE LAU LLE FCC TAT THT FLY ANA GAU GTC TTC GCC 511 ALB ANA HAA GAC GAC GAC GT TTC GCC GAC TTC GCC GAC TTC 511 ALB ANA ALB ALB ALB ATG MAS FLE LAU LLE		TGG Trp	CAG Gln	ACA Thr	GCA Ala	CAG Gln	GGC Gly	CTA Leu	TAC Tyr	TGG. Trp	GCC Ala	тас тут	GAT Asp	GGG Gly	TCA Ser	AAC Asn	CTC Leu				
321 Leu Ala Am Ser Oll Wal Am Am Ale Ap Phe Ala Arg Leu Phe Ala Leu Vel Am 1021 Got Too ACA GAC GOT OUT ATO THI TOO TOO ANG GAC AAA TOO GAD THE GAA THE TO 34 Ala Cay Min Ama Ala Gal Tile Phe Ser The Uses Gill Lays The Gal The Gal The Too 34 Ala Cay Min Ama Ala Gal Tile Phe Ser The Uses Gill Lays The Gal The Gal The Too 351 Phe Cay The Gal Wal Arg Ama Gal Tar Car Coo And Car Gal Ama Cot The Too Too Coo 351 Phe Cay Min Car Car Car And And Car Car Coo And Car Gal Ama Cot The Too Too Coo 351 Leu Gal Gal Car Cod Gal Gal Cad Car Car Coo And Car Gal Ama Cot The Too Too Coo 351 Leu Gal Car Cod Got Ami And And And Car Car Coo Ama Car Gal Ama Cot The Too Too Cod 351 Leu Gay Ala Fin Ama The Ama The Ama Ama Tin Can Car Cod Ama The Wal Arg Art Too 155 Cod Cod Cod Cod Got Cot The God Cod Cot Too Cod And Omo Cot Gal The Too Too 156 Leu Gal Car Cod Got Ama Car And And Cad Art Cod Art Too Cod Cot Cod Too 351 Leu Gay Ala The Ama The Ama The Ama Ama Tin Can And Omo Cot Gal The Too Too 157 Die Cod Cod Cod Cod Cot The God Cod Cot Too Cod And Car Cod Cod Too 158 Log Cod Cod Cod Got Cot The God Cod Cot Too Too Can Too Cod Cod Too Too 151 Cod Gad Cod Cod And The Too And Gad Cod Car Cod Cod Cod The Cod Car Too Cod Cod Cod Too 451 Ser Gal Gad Cod Cod Cod Too Cod Cod Cod Cod The Cod Car Too Cod Cod Cod 451 Ser Glu Cill Leu Ama Oly Val Ama Mag Jile Pice Ama Can Cod Cod Too Cod Cod 451 Ser Gill Cill Leu Ama Cill Yea Car The Cod Car Too Cod Cod Cod Too 451 Cill Leu Mag Cill Pice Car Art Too Cod Cod Cod Cod Cod Cod Cod Cod Cod C					TAC Tyr	AAC Asn	CAG Gln	ATC Ile	GTA Val	CGT Arg	COC Arg	ATC 11e	GCA Ala	GTT Val	ACG Thr						
341 Ala Cys The Asp Ala Gly Ile Phe Ser Trp Lys Glu Lys Trp Glu Phe Glu Phe Trp 1881 CCA CTA TCT GOT GTO CAL GAL GAL GDC GTO CCA GAL CTA GAL GAL CTT TO TO CTA GAL CTA GLO CTA CTT TO TO CTA GAL GAL GAL GAL GAL CTT CCA TT GAL CTT TO TO CTA GTO CTA GAL ASA ASA ASA JATA TYA PERA DAR HIS CJA ART CCA TTO CAL GAL GAL CTT CTA GAL CTA GAL GAL GAL GAL GAL GAL GAL GAL GAL GA		CTT Leu	GCC Ala	AAC ASTI	AGC Ser	GAA Glu	GTC Val	AAC Asn	AAT Asri	GCG Ala	GAT Asp	TTC Phe	OCC Ala	COC Arg	CTC Leu	TTC Phe	GCC Ala	CTC Leu	GTC Val	GAC Авр	GTC Val
361 PTC Leu BET GLY WEI ATS ARD AND LIV ATG PERO AND HIS GLY AND PTC PHE TTC LAU 1141. CTC GGT GCC CCC GCT ACT ALC ADC CAC CAC CAC TCC TCC CCT TTC CCA GTC 351. Leu GLY AE PTC ALL THE AND THC AND AND LIP FOR PHE HIS PEOP TO HIM HE TTC ALL 1201. CCA TCT GGT CAC GGO ACT TTT GGC GGT GGT GTT GTC AAA GTG GTC GGT GGT ATA TAC 1210. CCA TCT GGT CAC GGO ACT TTT GGC GGT GGT GTT GTC AAA GTG GTC GGT ATA TAC ATG 1211. CCG AG GGT CAC AGO AGC GTT GGT GGT GGT GTT GTC AAA TG TTC GAT ATG ATG 1211. CCG AG GGT CAC AGO GGT GAT GGT GGT GGT GTG GTC GGT ATG ATG ATG 1211. CCG GAG GGT CAAC CGG GGT AGC CGG GGA CAC ATT GGT CGC TAT GAT CGC AGO 1211. CCG GAG GGT CAAC CGG GGT AGC CGG GGA CAT GGT CGC CT TAT GGT CGC AGO 1211. CCG GAG GAC CGA AGO CGG GGT AGC CGG GGA CAT GGT CGC CGT TTC GGC CGC TTC GGT CGG 1211. CG GAG CGC AAC CGG GGT AGC CGG GGA CAT GGT CGC CGC TTTC GGC CGC TTC GGT CGG 1211. CG GAG CGC AGA CGG GGT AGC CGG CGC CGC ATG GTT CGC CGC TTC GGT CGG 1211. CG GAG GGC CGA AGC CGG GGT AGC CGG CGC CGC ATG GTC CGC CGC CGC GGG GGT 1211. CG GAG CGC CGG CGG GGT AGC CGG CGC CGC ATG GTT CGC CGC TTC GGT CGC 1211. CG GAG CGC CGC GGC GGT AGC CGG CGC CGC ATG GTT CGC CGC TTC GGT CGC 1211. CG GAG CGC CGC GGC GGT AGC CGG CGC CGC ATG GTT CGC CGC TTC GGT GGT CGC 1211. CG GAG CGC CGC CGC GGT AGC CGG CGC CGC AGG GGC CGC 1211. CG GAG CGC CGC CGC GGT AGC CGG CGC CGC CAGO GGC CGC CGC CGC GGC 1211. CG GAG CGC CGC CGC GGT AGC CGG CGC CGC CGC CGC GGC GGC 1211. CG GAG CGC CGC CGC GGT AGC CGG CGC CGC CGC CGC GGC GGC GGC 1211. CG GAG CGC CGC CGC GGT AGC CGG CGC CGC CGC GGC GGC GGC GGC GGC		GCT Ala	TGC Cys	ACA Thr	GAC Asp	GCT Ala	GGT Gly	ATC Ile	TTT Phe	TCC Ser	TGG Trp	AAG Lys	GAG Glu	AAA Lys	TGG Trp	GAG Glu	TTC Phe	GAA Glu	TTC Phe	TGG Trp	CGC Ārg
381. Leu Giy Ale Fro Ala The Ann The Ann Ang Tie Fro Phe Hug Fro Pic Phe Fro Ala 101. Exi Gir Got Got Got Get THT Got Got Get THT THT CAA And GHT GHT GHT THT THT GOT FOR SEGURISE ALL THE Phe GIY GIY ALL Wal Phe GHT ME VIA AT GHT THT THT THT 1361. GOZ GOT GHT GHT ACA THG ANG GAL THA GAA COL GHT ANT THE ANT THE ATT THT 1362. GOZ GOT GHT GHT ACA THG ANG GAL THA GAA COL GHT ANT ATT GHT ATT ATT 1363. GOZ GOT GHT GHT ACA THG ANG GAL THA GAA COL GHT ANT ATT GHT ATT ATT 1364. GOZ GOT GHT GHT ACA THG ANG GAL THA GAA COL GHT ATT ATT ATT 1364. GOZ GOT GHT GHT ACA THG ANG GAL THA GAA COL GHT ATT ATT ATT 1364. GOA GAL CHA ACT COL ANT COL GAL COL GAT THT GHT ATT ATT 1364. GAA GAL CHA CHT GHT THE LEVE AND AND THAT SITE VAL ATT GHT GHT ATT ATT 1364. GAA GAL CHA CHT GHT ATT COL GAL CHA CHT GHT GHT CHA CHT GHT GHT 1410. GAA CHA CHA CHT GHT ATT THO COL AFT GHT CHA CHA CHA CHA GHT THA 1411. GAA CHA CHT GHT GAL ALE GOL AFT THG CHA CHA CHA CHA CHA CHA CHA CHA 1511. GAL GHT GHT GAA ALE GOL AFT THG CHA CHA CHA CHA ANG GAL THG THAT SHT 1511. ATT ANA ANT ANG THE THE HUE HE FO THT THT FUNG ANG CHA CHA CHA CHA CHA CHA 1511. ATT ANA ANT ANG THE LEW LIFE FO THT THT FUNG ANG THA CHA CHA CHA GHT 1511. ATAL ANA ANT ANG THE LEW LIFE FO THT THT FUNG ANG THA CHA CHA CHA GHT GHT 1511. ATAL ANA CHA CHT OT THT CHA	1081 361	CCA Pro	CTA Leu	TCT Ser	GOT Gly	GTG Val	CGA Arg	GAC Asp	GAC Asp	GGC GIY	cot Arg	CCA Pro	GAC ASD	CAT His	GGA Gly	UAT Asp	CCT PTC	TTC Phe	tos trp	CTC Leu	ACT Thr
130. CCA TCT GOT CAC GOG ACC TTT GOC GOT GTT STC CAA ATG GTG CGT TGA TAC TAC S01 PFG Set GJy Kis Alla Thr Phe GJy GJy Alla Val Phe GJy Mek Val Arg Arg TYT TYT FYT 154. GOT CGC GTG GOT ACA TAC GAA GA CGC CAC ACA ATT GCA TC GAT GAT 411 GJy Arg Val GJy THT TTT Lys Asg Asg CGU PFG ARG CGT ACA ATT GCA TC GAT 412 GJY ATG Val GJY THT TTT Lys Asg Asg CGT A GGC CAL CTT TAC ATG CAC 413 Set CJU GL Hau And CJY Val Ann Arg Yag Lea Arg GIN FFT DYT ANG PFC ATG ATG 414 Set CJU GL Hau And CJY Val Ann Arg Yag Lea Arg GIN FFT DYT ANG PFC ATG CC 415 Set CJU GL Hau And CJY Val Ann Arg Yag Lea Arg GIN FFT DYT ANG PFC ATG CC 415 Set CJU GL Hau And CJY Val Ann Arg Yag Lea Arg GIN FFT DYT ANG PFC ATG CC 415 Set CJU GL Hau And CJY Val Ann Arg Yag Lea Arg GIN FFT DYT ANG PFC ATG CC 415 Set CJU GL Hau And CJY Val Ann Arg Yag Lea Arg GIN FFT DYT ANG PFC ATG CC 415 Set CJU GL Hau And CJY Val Ann Arg Yag Yag Lea Arg GIN FFT DYT ANG PFC ATG CC 415 Set CJU GL Hau And CJA ACC CC OT ATG CTG CC CC CT CT CC CT CT CC CT CT CC CT CT	1141 381	CTC Leu	GGT Gly	GCC Ala	CCA Pro	GCT Ala	ACT	AAC Asn	ACC Thr	AAC Ast	GAC Asp	ATT Ile	CCA Pro	TTC Phe	AAG Lys	CCT Pro	CCT Pro	TTC Phe	CCA Pro	GCT Ala	TAC TY:
1841 GOD COC OTA GOT ACA TGG AAG GAC GAC GAA COC GAC AAC ATT GC ATT GCA TG GAT ATG ATG 411 Giy Arg Val Giy THE TTE LOY ANY ANY ANY ANY ANY ANY ANY ANY ANY AN		CCA Pro	TCT Ser	GGT Gly	CAC	GCG Ala	ACC	TTT Phe	qqc gly	GOT GIY	GCT AIa	GTG Val	TTC Phe	CAA Gin	ATG Met	org Vai	CGT	CGA Arg	TAC	TAC Tyr	AAC Asn
441 See Diu Clis Leu Aen Diy Val Aen Arg Arg Leu Arg Clin Pro Tyr Arg Pro Thr Ala 1381 Art GAA GAC GAA CCC GOT Art GTO DOC ACC DOC Art GTT GOC GAC TTO GAC TOG GAC 451 INE GIU ME GIN Pro GIY Lie Val Arg Thr Arg Ti Le Yi Arg Marg Ber Ala 1441 GAA CTC ATG THO GAA AAC GOC ATT TOG GOC ATC TTC CTC GTC GTC GAC TOG CGT TTC 451 GI Leu Mee The GUN ARA ALG GOC ATT TOG GOC ATC TTC CTC GTC GTC GAC TOG CGT TTC 451 GI Leu Mee The GUN ARA ALG SEC ATT TOG GOC ATC TTC TTC GTC GTC GAC TOG CGT TTC 551 Ala Ala Ala Ala Ala Arg Arg Tie Hee Ser Arg Tie Thr Thr Lys Arg GAC GTC GTC 551 Ala Ala Ala Ala Ala Arg Arg Tie Leu Tie Pro Thr Thr Lys Arg Cac Tog CGT 551 Ana CAC GOC GCC GAC GTT TTC GUN ACC GTA GGAC ATT AGG TAC GAC ACC ACC AGO 552 Ana Arg Tog Leo Acc GTO TTC CAU ANC GTA GGAC ATT AGG TAC GAC ACC AGO CAC 551 Ana Ard GOC GCC GAC GTO TTC CAU ACC GTA GGAC ATT AGG TAC GAC ACC AGO CAC 551 Ana Ard TOg Leo Acc GTO TTC CAU ANC GTA GGAC ATT AGG TAC GAC ACC AGO CAC 551 Ana Ard TOg Leo Acc GTO TTC CAU ANG GTA GGAC ATT AGG TAC GAC ACC AGO CAC 551 Ana Ard TOg Leo Acc GTO TTC CAU ANG GTA GGAC ATT AGG TAC GAC ACC AGO CAC 551 Ana Arg Tog Leo TTC CAU ANG GTA GGAC GTA GAC GAT TA GGA GAC GTA THA ANG TAC GTA GTA GGAC GTAG	1261	0GC	CGC	GTA	GGT	ACA	TGG	AAG	GAC	GÁC	GAA	000	GAC	AAC	ATT	GCC	ATC	GAT	ATG	ATG	ATC
1981 AFC GAA GAC CAA CCC GOT AFC GTG DOC ACC DOC AFT GTT GOT GAC TOG GAC TOG GAC 461 ILG GLU AFG GIN FFG GIY ILE VAL AFG THE VAT AFG YLE VAL AFG KLE FAR AFG SER XIA 141 GAA CTC AFG TFC GAA AAC GCC AFT TOG GOC AFC TFC CTC GOT GTC CAC TOG COT 463 GLU GLU AFG FG GUL AFG ATA XIE ILE FFG FTG FTG CTC GAT GTC CAC TOG COT 464 ILG XIE AFG FG GUL AFG XIE ILE FFG THT FTT FUX AFG VAL TYF AFG FAR 1501 ACC AGC GOC GOC GAC AFT CTC AFT GOC AGC AFC AAG GAC GTC TAC GOT GTC 501 ANIA XIE AHA XIE AFG AFG ILE FEW ILE FFG THT FTT FUX AFG VAL TYF AIG VAL 1561 AAC AAT GOC GOC GOC GTT FC GAU GAC GTT GAG GAC ATT AGG TAC AFC AGC GOC 511 ANIA AFT GIY AIB TTT CAU GAAC GTT GAG GAC ATT AGG TAC AFC AGC GOC 511 ANIA AFT GIY AIB TTT VAL FFW GIN GAT VAL GUL AFT ILE AFG TYF TTT FLE AFG GUY		7CG Ser	GAG Glu	GAS Glu	CTC Leu	AAC	GGC Gly	org Val	AAC Asn	CGC	GAC ASD	OTA Leu	CGC Arg	CAG Gin	CCT	TAT	GAC ASE	DCC Pro	ACG Thr	GCC Ala	CCA Pro
1441 GAA CTC ATS TTC GAA AAC GCC ATT TOG CGC ATC TTC GTC GTT GTC CAC TGG GTT TTC 481 GIL Leu Nec Phe GIL ANN AIS ILE SET ATG ILE Phe Leu GIY Val KIS TTP ATG Phe 1501 GGC SCC GCC GCC GGC GAA ATC CTC ATT CCC AGO AGO ACA AAG GAG GTC TAC GGT GTC 501 AIS Als Als Als ATG AND ILE Leu ILE PTC THT THT THT Lys AND Yal Tyr Als Val 1561 AAC AAT GGG GCC ACC GTT TTC CAT AAC GTA GTG GTAC TAT AGO TAT CTC 521 AAN ATG GIY Als THT VAl Phe GIN AND YAI GIL AND ILE ATG TYr THT THE ACG GIY		ATC	GAA	GAC	CAA	000	GGT	ATC	GTG	000	ACC	CGC	ATT	GTT	CGC	CAC	TTC	GAC	TCG	GCC	TGG
1501 Get dec dec dec and art etc att etc and act act aca and gac get tat get att 501 Alls Alls Alls Alls Arty Arg Ise Lew lie Peo The The Type Arg Val Type Alls Val 1561 AAC ART GOU GEC ACC GTO THE CASI AND GTA GAC ACT ACG TAC ACA ACC GO GOT 521 Ant Arg Cly Alls The Val Phe Clin Arg Val Cliv Arg Ise Arg Type The The Acg OLY	1441	GAA	orc	ATG	TTO	GAN	AAC	ggg	ATT	TOG	CGC	ATC	TTC	CTC	sor	ore	CAC	TOG	CGT	TTC	GAT
1561 AAC AAT GOC GOC ACC GTG TTC CAU AAC GTA GAG GAC ATT AGG TAC ACA ACC AGG GGT 521 ABN ABN GUy Ala Thr Wal Phe Gin Abn Val Giu Abn Tie Arg Tyr Thr Thr Arg Giy	1501																				
and the out and the low low the control the are adding into any low the rest	1561	AAC	TAA	GGC	800	ACC	070	TTC	CAG	AAC	OTA	GAG	GAC	ATT	AGG						
541 Ard Glu Ast Pro Glu Gly Leu Phe Pro 11e Gly Gly Val Pro Leu Gly Tie Glu Tie																					Thr
(t)				10																	
561 Asp Giu Ile Phe Asn Ash Gly Leu Lys Pro Thr Pro Pro Glu Ile Gin Pro Met Pro	561	Asp	Glu	Ile	Phe	Asn	Ast	Gly	1eu	Lys	Pro	Thr	Pro	Pro	Glu	lle	Gin	Pro	Met	Pro	Gin
1741 GAG ACA CCG GTG CAG AAG CCG GTG GGA CAG CCG GTT AAG GGC ATG TGG GAG GAA 581 Glu Thr Pro Val Gin Lys Pro Val Giy Gin Gin Pro Val Lys Giy Met Trp Giu Glu	581	Glu	Thr	Pro	Val	Gln	Lys	Pro	Val	dīy	Glh	Gln	Pro	Va.	Lys	Gly	Met	Trp	Glu	Glu	glu
1861 DAG GCG CCG OTA GTC AAG GAG GCG CCG TAG atgaggtatggggtaggggtgaggg 601 Gln Ala Pro Val Val Lys Glu Ala Pro AMB	1801 601	CAG Gln	GCG Ala	CCG Pro	GTA Val	GTC Val	AAG Lys	GAG Glu	GCG Ala	CCG Pro	TAG	atga	ggta	togt	gaga	gtat	aaaa	tagg	cgtg	agtg	gga
cgtcttggagtaagggctgatgoggaagtttagtttgattggtacgggtctggtttagggcgtgogctcgatactctgcggt																	ates	atao	LCLS	leggt	taa
tacacctatttegatattagatagaatgeteagatacetagaatgtgatgatgatteagetgtet																					

Fig. 2. Nucleotide sequence for the vanadium chloroperoxidase cDNA from *C. inaequalis*. The transcription initiation site is indicated by a solid arrow (\rightarrow) . The sequence given in italics was only obtained from the sequence analyses of genomic DNA while the rest of the sequence was obtained from sequence analyses of cDNA. The complete 3' region up to base pair number 180 and the 5' region from base pair number 1141 to the *pvull* site (cagetg) located at the end of the transcribed region, were further confirmed by double-stranded sequencing of a genomic DNA clone containing the gene encoding the vanadium chloroperoxidase. A typical TATAAA box upstream of the transcription initiation site is underlined. The amino acid sequence of the predicted translation product is shown below the nucleotide sequence. The vanadium chloroperoxidase coding region is flanked by a 75-bp 5' noncoding region and a 186-bp [excluding the poly(A) tail] 3' noncoding region. No leader peptide can be assigned to the extracellular enzyme. Potential N-glycosylation sites at positions 128 and 274 are marked (***). A putative polyadenylation signal is overlined. Primers used for amplification reactions are indicated (\rightarrow). Vertical crosses (\uparrow) indicate the position of an Asp-Pro bond (see text).

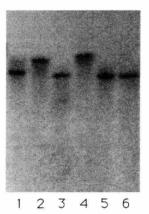


Fig.3. Southern-blot analysis of genomic DNA from *C. inaequalis* probed with a RT-PCR fragment for the vanadium chloroperoxidase. The genomic DNA from *C. inaequalis* was digested with different restriction enzymes and subsequently separated by electophoresis in a 0.5% agarose gel, transferred to a nitrocellulose membrane, and hybridized with a "P-labelled probe as described in the text. Lanes 1–6 contain genomic DNA digested with *Eco*RI, *Bam*HI, *Eco*RI/*Bam*HI, *XbaI*, *Ps*f and *XbaI/Pstl*, respectively.

67488 Da, which is in good agreement with the experimentally determined molecular mass of about 67 kDa, as found by SDS/ PAGE according to Laemmli [20]. Although the protein was extracellular, no putative leader peptide, normally present in precursors of extracellular proteins [28, 29] can be assigned. In order to ascertain whether the first encoded methionine residue was present in the mature protein, the protein was treated with CNBr in order to remove this putative N-terminal methionine residue. The resulting peptide mixture was submitted to amino acid sequence analysis. No amino acid sequence in the N-terminal region could be detected indicating that the mature protein lacks a N-terminal methionine residue. A tryptic digestion of the chloroperoxidase yielded a sequenced amino peptide corresponding to residues 12–19 of the encoded protein (Table 1).

Liu et al. [18] reported that the enzyme is tetrameric. This is in agreement with our results presented in Fig. 1, which show two active bands on SDS/PAGE gels, one in the high-molecularmass range, being a multimeric form of the vanadium chloroperoxidase, and one with a lower molecular mass, representing a monomer. The multimeric form is apparently stable as it is not affected by the presence of SDS (Fig. 1).

It was previously suggested for the vanadium bromoperoxidase from the seaweed A. nodosum that the thermostability and chemostability of the enzyme are (at least partly) due to the presence of eight disulphide bridges [30]. Therefore, it was investigated if the remarkable stability of the vanadium chloroperoxidase could be explained by the presence of disulphide bridges. Determination of thiols in the protein by the method of Thannhauser et al. [24] revealed the presence of two reactive thiols. As the deduced amino acid sequence shows the presence of only two cysteine residues (Fig. 2), these results indicate that no disulphide bridges are present in the mature protein. Treatment of the protein with 2-mercaptoethanol did not affect the mobility of the protein when compared to the non-treated sample, which is in agreement with the finding that the protein lacks disulphide bridges. An intriguing phenomenon occurred when the vanadium chloroperoxidase was submitted to denaturing SDS/PAGE. Minor bands appeared in the gel, which increased in intensity upon prolonged boiling of the sample (data not shown). Subsequent amino acid sequencing of the resulting peptides showed that the protein is cleaved at Asp-Pro sequences. The vanadium chloroperoxidase contains six of these Asp-Pro sequences (Fig. 2). Interestingly, a similar cleavage pattern was observed for vanadium chloroperoxidase isolated from the fungus *D. biseptata* (unpublished results), after submitting the sample to denaturing SDS/PAGE. This finding supports the suggestion that the fungal vanadium chloroperoxidases are homologous.

Metal content of the vanadium chloroperoxidase. As reported previously, the vanadium content is 1 mol vanadium/mol chloroperoxidase [14]. According to Liu et al. [18], the enzyme contained 0.7 mol iron and 2.2 mol zinc as deduced from X-ray fluorescence. However, the role of these metals in enzyme activity was not established. In our experiments, zinc and/or iron addition had no effect on the enzyme activity. We determined the amount of zinc, iron, and also calcium and copper, using purified vanadium chloroperoxidase. The measurements were carried out on four different batches of isolated chloroperoxidase. After dialysis for 24 h in 25 mM Tris/SO4, pH 8.5, the samples were passed through a Sephadex G-25 or Chelex-100 column to remove all residual traces of metal ions. After both procedures, vanadium was found in a ratio of 1 ± 0.1 mol/mol enzyme, as expected. Zinc was found in a ratio of $0.3 \pm 0.04/$ mol protein. Addition of 2 mM 1,10-phenanthroline or 10 mM EDTA to the dialysis buffer did not affect the amount of zinc present in the protein, indicating that the zinc is strongly bound. Iron and copper were found at ratios less then 0.2 and 0.02 mol/ mol enzyme, respectively. Although calcium was found in a ratio of about 1 mol/mol enzyme after passing it through a G-25 column, this ratio decreased to less then 0.2 after passing the sample through a Chelex column.

Glycosylation of the vanadium chloroperoxidase. After treatment of the vanadium chloroperoxidase with N-glycanase, the mobility of the treated enzyme was not altered when compared with the untreated enzyme. This strongly indicates that no Nlinked sugars are present in the purified protein, although two putative N-glycosylation sites are present in the deduced primary structure (Fig. 2). A general glycan detection kit was also used to test if the protein was glycosylated. No reaction was observed indicating that the protein is not glycosylated.

Possible role of methionine in the halogenation catalysis. Comparison of the primary structure of the vanadium chloroperoxidase from C. inaequalis with bacterial non-haem haloperoxidases from Pseudomonas pyrrocinia [31]. Streptomyces aureofaciens [32, 33], and Streptomyces lividans [34] showed that there is one small region region sharing amino acid sequence similarity (Fig. 4). Since these bacterial haloperoxidases did not contain metals, amino acid residues were proposed to be involved in catalysis of the halogenation reaction [35]. For the bromoperoxidases of S. aureofaciens (BPO-A1 and BPO-A2), it was shown that incubation of the enzymes in the presence of cyanide, bromide, and hydrogen peroxide resulted in cleavage of the enzymes after a specific methionine due to the formation of CNBr. Therefore, Haag et al. [35] proposed that this methionine residue is involved in the catalytic activity of the bromoperoxidases. This methionine residue is located in the region showing similarity with the vanadium chloroperoxidase from C. inaequalis (Fig. 4). To study if a similar cyanide-dependent cleavage

CPO	c.	inaequalis	DAVRVAI	AMGGAQALN STGGGEVAR	STKRSP280
CPO	5.	lividans	GALVHIGH	SMGGGELAR	Y V A R A E
		aureofaciens	DVTLVAH	SMGGGELAR	Y V G R H G 108
		aureofaciens	DAVLVGF	SMGIGEVAR	YVSSYG
		pyrrocinia	NALVINIGH	STGGGEVAR	YVANDG
		sp. KKS102	KAHLVGN	SMGGAGALN SFGGGIALA	FALEYPIZE
DmpD	Ρ.	sp. CF600	QADLVGN	SFGGGGIALA	LAIRHP121
			4 M 4 M	19. M. M. H.	* *

Fig. 4. Sequence comparison of vanadium chloroperoxidase from C. inaequalis with bacterial non-haem haloperoxidases and two other adfi-hydrolase-fold enzymes. The figure shows the alignment of one small region of the vanadium chloroperoxidase sequence sharing similarity with six alfi-hydrolase-fold enzymes from S. lividans chloroperoxidase [34], S. aureofaciens bromoperoxidases A1 and A2 [32, 33, 36], P. pyrrocinia chloroperoxidase [31], Pseudomonas sp. strain KKS102 BphD (2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase) [37], and Pseudomonas CF600 DmpD (2-hydroxymuconate-semialdehyde hydrolase) [38], Boxed regions indicate identical residues. Where five or more sequences, including the CPO from C. inaequalis, contain the same residue an asterisk (*) appears below the residue; a plus symbol (+) denotes three or four sequences, including the CPO from C. inaequalis with the same residue.

occurs in the fungal vanadium chloroperoxidase, the enzyme was incubated with cyanide, bromide, and hydrogen peroxide, as described for the bacterial haloperoxidases [35], and subjected to SDS/PAGE. No cleavage was observed suggesting that this methionine residue is not involved in the catalytic activity of the fungal vanadium chloroperoxidase. This region also has sequence similarity with other α/β -hydrolase-fold enzymes, a group with an increasing number of members to which the bacterial non-haem haloperoxidases now belong [36], though the potentially functional nucleophile is not conserved in the chloroperoxidase sequence. The significance of this similarity, in other than a purely structural sense is not clear, though it is interesting to note that the other aligned hydrolase sequences (Fig. 4) are involved in a dehalogenation pathway [37–39].

DISCUSSION

The results described in this study show that the vanadium chloroperoxidase of *C. inaequalis* is not a normal secretion protein. A signal sequence to guide a precursor into the endoplasmic reticulum, as found for many eukaryotic extracellular proteins, is lacking. The finding that the vanadium enzyme might be secreted via an alternative route is supported by the observation that the vanadium enzyme is not N-glycosylated and does not contain disulphide bridges. These post-translational modifications normally take place in the endoplasmic reticulum. It is not clear how the vanadium enzyme is secreted by the fungus, but it would be interesting to study the mechanism by which the enzyme is crossing the membrane.

Although the N-terminal amino acid residue could not be determined by amino acid sequencing, the results described in this study strongly suggest that the mature vanadium chloroperoxidase does not contain methionine as a N-terminal amino acid residue but the first amino acid should be in close vicinity to the first encoded methionine residue. X-ray crystal structure analysis of the vanadium chloroperoxidase, also indicates that the first amino acid residue of the mature protein is in close vicinity to the first encoded methionine residue (Messerschmidt, A., unpublished results).

Another interesting observation is that the vanadium chloroperoxidase from C. *inaequalis* is secreted in an apoform when grown in the absence of vanadate [14]. This apoenzyme can be easily activated by the addition of vanadate. Apparently these enzymes do not require an additional enzyme system for the incorporation of the vanadate, but are able to obtain their prosthetic group directly from the environment.

A computer database (EMBL) was screened using the Blast program for the presence of proteins that share amino acid sequence similarity to the vanadium chloroperoxidase from C. inaequalis. No proteins having sequence similarity were found in this search. Several fungi belonging to the hyphomycetes are able to produce an apparently similar vanadium chloroperoxidase (unpublished results). It therefore appears that the vanadium chloroperoxidases constitute a new family of peroxidases present in the terrestrial environment. Protein fingerprinting and immunoblotting experiments revealed that the fungal vanadium chloroperoxidases from C. inaequalis and D. biseptata are very similar to each other (unpublished results). However, no similarity was found with peptide sequences of the vanadium bromoperoxidase from the seaweed A. nodosum [40]. Therefore, it is presently unclear whether the fungal haloperoxidases are structurally related to the haloperoxidases from marine sources. EPR spectra from both fungal vanadium chloroperoxidases and the vanadium bromoperoxidase from A. nodosum [14], however, are very similar, indicating that the environment of the vanadate is identical in these haloperoxidases.

It was demonstrated for three bacterial non-haem haloperoxidases that methionine residues are involved in the halogenation reaction [35]. It has been suggested for these enzymes that the reaction mechanism involves the oxidation of methionine to methionine sulphoxide, which subsequently reacts with halide ions [35]. Our results indicate that in the vanadium chloroperoxidase from *C. inaequalis*, a methionine is not involved in catalysis. Several experiments on the vanadium bromoperoxidase from the seaweed *A. nodosum* indicate that the vanadium in this enzyme is not involved in a redox reaction but serves to bind the hydrogen peroxide and to activate it to allow reaction with bromide and subsequent formation of HOBr [12, 13]. The vanadium in the fungal vanadium chloroperoxidase probably has a similar function in the halogenation reaction and on the basis of steadystate kinetics a peroxo intermediate has been postulated [17].

Although the fungal vanadium chloroperoxidases were found to be very thermostable and chemostable, no disulphide bridges are present in the protein. Secondary structure prediction, using the Garnier prediction method (PC-Gene), indicate that the protein mainly consists of *a*-helix structures (40-50%), which could be involved in stabilization of the vanadium enzyme. The metal analysis of the vanadium chloroperoxidase shows that iron, copper, and calcium do not constitute part of the enzyme. The amount of zinc is also apparently too low to constitute part of the monomer of the vanadium chloroperoxidase monomer and the determined zinc content of about 0.3 mol/ mol enzyme could be due to a high background of zinc that is hard to remove.

There are six Asp-Pro bonds, which are known to be weak peptide bonds [41], found in the protein, which easily break upon boiling of the protein sample. This is especially striking as the protein was found to be very stable at temperatures up to 80 °C or in the presence of different organic solutions [17]. The Asp-Pro sequences might play an essential role in folding of the protein possibly providing sharp turns in the tertiary structure.

The expression studies described in this study clearly show that the vanadium enzyme is produced in the secondary growth phase, the so-called idiophase. The expression is regulated at the level of transcription and is apparently dependent on nutrient limitations. Previously it has been suggested that the fungal vanadium chloroperoxidases might be involved in degradation and/ or penetration of plant tissues (unpublished results). This would be in agreement with the expression of the *C. inaequalis* vanadium chloroperoxidase in the secondary growth phase as a result of nutrient depletion. The mechanism by which the vanadium enzyme is involved in plant tissue disruption is unclear. It is conceivable that the hypochlorous acid formed by the enzyme directly oxidizes the polymeric structures present in the plant cell wall in order to solubilize these rather tough structures before penetration. It has been suggested for cellulolytic and lignolytic micro-organisms that a small oxidizing agent, e.g. hydrogen peroxide, is necessary to initiate the solubilization of the polymeric backbone of the plant tissues [42, 43], before cellulases and/or ligninases take over the breakdown of these polymeric structures.

The authors wish to thank the Department of Molecular Microbiology of the Free University, Amsterdam, for the use of their DNA Sequencer and their help in operating this system. This work was supported by grants from the Netherlands Foundation for Chemical Research (SON) and was made possible by financial support from the Netherlands Organization for Scientific Research (NWO) and the Netherlands Technology Foundation (STW). We thank Dr L. B. J. Vertegaal (SON/STW) for his stimulating role in the organization and sponsoring of this project.

REFERENCES

- Wever, R. & Kustin, K. (1990) Vanadium: a biologically relevant element, Adv. Inorg. Chem. 35, 81-115.
- Butler, A. & Walker, J. V. (1993) Marine haloperoxidases, *Chem. Rev.* 93, 1937–1944.
- Neidleman, S. L. & Geigert, J. (1986) Biohalogenation: principles, basic roles and applications, Ellis Horwood Ltd., Chichester.
- Gribble, G. W. (1992) Naturally occurring organohalogen compounds – a survey, J. Nat. Products 55, 1353–1395.
- Hager, L. P., Morris, D. R., Brown, F. S. & Eberwein, H. (1966) Chloroperoxidase. II. Utilization of halogen anions, J. Biol. Chem. 241, 1769-1777.
- Vilter, H. (1994) Peroxidases from Phaeophycae: a vanadium(V) dependent peroxidase from Ascophyllum nodosum, Phytochemistry (Oxf.) 23, 1387-1390.
- De Boer, E., Van Kooyk, Y., Tromp, M. G. M., Plat, H. & Wever, R. (1986) Bromoperoxidase from Ascophyllum nodosum: a novel class of enzymes containing vanadium as a prosthetic group? *Biochim. Biophys. Acta 869*, 48 – 53.
- Soedjak, H. S. & Butler, A. (1991) Mechanism of dioxygen formation catalyzed by vanadium bromoperoxidase from *Macrocystis pyrifera* and *Fucus distichus*: steady state kinetic analysis and comparison to the mechanism of V-BrPO from *Ascophyllum nodosum*, *Biochim, Biophys. Acta 1079*, 1–7.
- Plat, H., Krenn, B. E. & Wever, R. (1987) The bromoperoxidase from the lichen *Xantoria parietina* is a novel vanadium enzyme, *Biochem. J.* 248, 277-279.
- De Boer, E., Boon, K. & Wever, R. (1988) Electron paramagnetic resonance studies on conformational states and metal ion exchange properties of vanadium bromoperoxidase, *Biochemistry* 27, 1629–1635.
- Arber, J. M., De Boer, E., Garner, C. D., Hasnain, S. S. & Wever, R. (1989) Vanadium K-edge X-ray absorption spectroscopy of bromoperoxidase from Ascophyllum nodosum, Biochemistry 28, 7986-7973.
- Kusthardt, U., Hedman, B., Hodgson, K. O., Hahn, R. & Vilter, H. (1993) High-resolution XANES studies on vanadium-containing haloperoxidase: pH-dependence and substrate binding. *FEBS Lett.* 329, 5–8.
- Wever, R., Krenn, B. E., De Boer, E., Offenberg, H. & Plat, H. (1988) Structure and function of vanadium-containing bromoperoxidases, *Progr. Clin. Biol. Res.* 274, 477–493.
- Van Schijndel, J. W. P. M., Vollenbroek, E. G. M. & Wever, R. (1993) The chloroperoxidase from the fungus *Curvularia inae-qualis*; a novel vanadium enzyme, *Biochim. Biophys. Acta 1161*, 249–256.
- Asplund, G. & Grimvall, A. (1991) Organohalogens in nature; more widespread than previously assumed, *Environ. Sci. Technol.* 25, 1347-1350.

- Van Schijndel, J. W. P. M., Simons, L. H., Vollenbroek, E. G. M. & Wever, R. (1993) The vanadium chloroperoxidase from the fungus *Curvularia inaequalis*; evidence for the involvement of a histidine residue in the binding of vanadate, *FEBS Lett.* 336, 239– 242.
- Van Schijndel, J. W. P. M., Barnett, P., Roelse, J., Vollenbroek, E. G. M. & Wever, R. (1994) The stability and steady-state kinetics of vanadium chloroperoxidase from the fungus Curvularia inaeaudis, Eur. J. Biochem. 225, 151-157.
- Liu, T. E., M'Timkulu, T., Geigert, J., Wolf, R., Neidleman, S. L., Silva, D. & Hunter-Cevera, J. C. (1987) Isolation and characterization of a novel non-heme chloroperoxidase, *Biochem. Biophys. Res. Commun.* 142, 329–333.
- Gross, E. (1967) The cyanogen bromide reaction, Methods Enzymol. 11, 238-255.
- Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* 227, 680-685.
- Schägger, H. & Von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. 166, 368-379.
- Matsudaira, P. (1987) Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes, J. Biol. Chem. 262, 10035-10038.
- 23. De Boer, E., Plat, H., Tromp, M. G. M., Wever, R., Franssen, M. C. R., Van der Plas, H. C., Meijer, E. M. & Schoenmaker, H. E. (1987) Vanadium containing bromoperoxidase: an example of an oxidoreductase with high operational stability in aqueous and organic media, *Biotechnol. Bioeng.* 30, 607-610.
- Thannhauser, T. W., Konishi, Y. & Scheraga, H. A. (1987) Analysis for disulphide bonds in peptides and proteins. *Methods Enzymol.* 143, 115-119.
- Sambrook, J., Fritch, E. F. & Maniatis, T. (1989) Molecular cloning: a laboratory manual, 2nd edn, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Siebert, P. D. & Larrick, J. W. (1992) Competitive PCR, *Nature 359*, 557–558.
- Proudfoot, N. J. & Brownlee, G. G. (1976) 3' Non-coding region sequences in eukaryotic messenger RNA, *Nature* 263, 211–214.
- Perlman, D. & Halvorson, H. O. (1983) A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J. Mol. Biol. 167, 391–409.
- Von Heijne, G. (1985) Signal sequences: the limits of variation, J. Mol. Biol. 184, 99-105.
- Tromp, M. G. M., Olafsson, G., Krenn, B. E. & Wever, R. (1989) Some structural aspects of vanadium bromoperoxidase from Ascophyllum nodosum, Biochim. Biophys. Acta 1040, 192–198.
- Wolfframm, C., Lingens, F., Mutzel, R. & Van Pée, K.-H. (1993) Chloroperoxidase-encoding gene from *Pseudomonas pyrrocinia*, *Gene (Amst.) 130*, 131–135.
- Pfeifer, O., Pelletier, I., Altenbuchner, J. & Van Pée, K.-H. (1992) Molecular cloning and sequencing of a non-haem bromoperoxidase gene from *Streptomyces aureofaciens* ATCC 10762, *J. Gen. Microbiol.* 138, 1123–1131.
- Pelletier, I., Pfeifer, O., Altenbuchner, J. & Van Pée, K.-H. (1994) Cloning of a second non-haem bromoperoxidase gene from *Strep tomyces aureofaciens* ATCC 10762: sequence analysis, expression in *Streptomyces lividans* and enzyme purification. *Microbiology* 140, 509–516.
- Bantleon, R., Altenbuchner, J. & Van Pée, K.-H. (1994) Chloroperoxidase from *Streptomyces lividans*: isolation and characterization of the enzyme and the corresponding gene, *J. Bacteriol.* 176, 2339–2347.
- Haag, T., Lingens, F. & Van Pée, K.-H. (1991) A metal ion and cofactor-independent enzymic redox reaction: halogenation by bacterial non-heme haloperoxidases. *Angew. Chem. Int. Ed. Engl.* 30, 1487–1488.
- Hecht, H. J., Sobek, H., Haag, T., Pfeifer, O. & Van Pée, K. H. (1994) The metal-ion-free oxidoreductase from *Streptomyces aureofaciens* has an *alfi* hydrolase fold, *Nature Struct. Biol.* 1, 532-537.
- Kimbara, K., Hashimoto, T., Fukuda, M., Koana, T., Takagi, M., Oishi, M. & Yano, K. (1989) Cloning and sequencing of two

tandem genes involved in degradation of 2,3-dihydroxybiphenyl to benzoic acid in the polychlorinated biphenyl-degrading soil bacterium *Pseudomonas* sp. strain KKS102, *J. Bacteriol.* 171, 2740-2747.

- Nordlund, I. & Shingler, S. (1990) Nucleotide sequences of the meta-cleavage pathway enzymes 2-hydroxymuconic semialdehyde dehydrogenase and 2-hydroxymuconic semialdehyde hydrolase from *Pseudomonas* CF600, *Biochim. Biophys. Acta 1049*, 227–230.
- Janssen, D. B., Pries, F. & Van der Ploeg, J. (1994) Genetics and biochemistry of dehalogenating enzymes, *Annu. Rev. Microbiol.* 48, 163-191.
- Tromp, M. (1992) Studies on vanadium-bromoperoxidases from seaweeds, Ph. D. thesis, University of Amsterdam, Amsterdam.
- Allen, G. (1981) Specific cleavage of the protein in Sequencing of proteins and peptides (Work, T. S. & Burdon, R. H., eds) vol. 9, pp. 50-51, Elsevier Science Publishers B. V., Amsterdam.
- Veness, R. G. & Evans, C. S. (1989) The role of hydrogen peroxide in the degradation of crystalline cellulose by Basidiomycete fungi, J. Gen. Microbiol. 135, 2799-2806.
- Evans, C. S., Gallagher, I. M., Atkey, P. T. & Wood, D. A. (1991) Localization of degradative enzymes in white-rot decay of lignocellulose, *Biodegradation* 2, 93–106.

Copyright of European Journal of Biochemistry is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.