Cloning and expression in *Escherichia coli* of the *nahA* gene from *Porphyromonas gingivalis* indicates that β -N-acetylhexosaminidase is an outer-membrane-associated lipoprotein

A. Lovatt and I. S. Roberts

Author for correspondence: I. S. Roberts. Tel: +44 533 522956. Fax: +44 533 523013.

Department of Microbiology and Immunology, Medical Sciences Building, P.O. Box 138, University of Leicester, University Road, Leicester LE1 9HN, UK

***Porphyromonas gingivalis* has been implicated in human periodontal diseases. It expresses a number of exoglycosidase enzymes capable of hydrolysing host proteoglycan residues. As a first stage to explore the role of these enzymes in periodontal tissue damage, the *nahA* gene of *P. gingivalis* W83, which encodes β -N-acetylhexosaminidase (β -Nahase), was cloned. The gene was expressed poorly in *Escherichia coli*, but increased expression was achieved by cloning the *nahA* gene downstream of the *tac* promoter. Southern blot analysis revealed that *nahA* was present as a single copy, and it was found in all the other *P. gingivalis* strains tested. In contrast, sequences homologous to *nahA* were not detected in either *P. endodontalis* or *P. asaccharolytica*. The *nahA* gene was 2331 bp long and encoded a β -Nahase enzyme of 777 amino acids with a predicted molecular mass of 87 kDa. A characteristic signal peptide for an acylated lipoprotein was present at the amino-terminus, suggesting that the mature β -Nahase is a lipoprotein. The predicted amino acid sequence of the *P. gingivalis* β -Nahase shared homology with the catalytic domains of the human β -Nahase enzyme and the chitinase of *Vibrio harveyi*, suggesting a common catalytic mechanism.**

Keywords: *Porphyromonas gingivalis*, β -N-acetylhexosaminidase, exoglycosidase, lipoprotein

INTRODUCTION

The development of periodontal disease is thought to be associated with the presence of several Gram-negative eubacteria in the subgingival region (Slots & Genco, 1984; Takazoe *et al.*, 1984). Within this group, *Porphyromonas gingivalis*, a Gram-negative obligate anaerobe, has been isolated in proportionally large numbers from the lesions of advanced adult periodontitis (Slots *et al.*, 1986; White & Mayrand, 1981). A variety of putative virulence factors have been described for *P. gingivalis*, including fimbriae, proteinases and collagenases, and a number of exoglycosidases (Holt & Bramanti, 1991; Mayrand & Holt, 1988). It has been suggested that the secretion of proteinases may play a number of important

roles; these include attachment to host tissues and to other oral micro-organisms (Li *et al.*, 1991; Nishikita *et al.*, 1989), degradation of molecules of the host immune system (Sundqvist *et al.*, 1985), and generation of oligopeptides and amino acids for the growth of *P. gingivalis* (Shah & Gharbia, 1989). A number of genes encoding for proteinases (Bourgeau *et al.*, 1992; Park & McBride, 1993) and a collagenase (Kato *et al.*, 1992), have been cloned and are currently the subject of detailed analysis.

Less well studied are the exoglycosidases, neuraminidase and β -N-acetylhexosaminidase (β -Nahase) EC 3.2.1.52, both of which are secreted by *P. gingivalis* (Greenman & Minhas, 1989; Mayrand & Holt, 1988; Tipler & Embery, 1985). Exoglycosidases catalyse hydrolytic cleavage of the glycosidic linkages beginning at the outer, non-reducing, end of each oligosaccharide chain. The neuraminidase cleaves sialic acid residues occurring at the termini of both simple and complex oligosaccharides and on glycoproteins, whilst the β -Nahase will remove terminal N-acetylglucosamine and

Abbreviations: GAG, glycosaminoglycan; MUAG, 4-methylumbelliferyl-N-acetyl-D-glucosaminide; β -Nahase; β -N-acetyl-hexosaminidase.

The GenBank accession number for the nucleotide sequence reported in this paper is X78979.

N-acetylgalactosamine residues (Cabezas, 1989). These enzymes often act sequentially to degrade complex oligosaccharides, and as a result may increase the sensitivity of glycoproteins to subsequent proteolytic cleavage (Jensen & Ledet, 1986).

Proteoglycans are major constituents of gingival connective tissue and are made up of proteins covalently linked to a large number of glycosaminoglycan (GAG) chains and oligosaccharides, the principal GAG in gingival tissue being dermatan sulphate, with heparan sulphate and chondroitin sulphate present as minor components (Bartold *et al.*, 1982; Bartold, 1987). Proteoglycans function in regulating cell adhesion and growth, matrix formation, collagen fibril formation and binding of growth factors (Rouslahti, 1989; Uitto, 1991). Clearly, disruption of proteoglycan function by breakdown of GAG by exoglycosidases may have pronounced effects on the functional integrity of gingival tissue. Indeed, a number of microbial β -Nahases have been implicated in the disruption of the host intestinal mucosal surface, by the degradation of mucins and proteoglycans, which allows the subsequent invasion of the underlying host tissue (Boureau *et al.*, 1993; Werries *et al.*, 1983). In addition, the ability of β -Nahase to remove glycosidic residues from glycoproteins such as IgG (Koide *et al.*, 1977) will have significant effects on the function of these molecules and the possible outcome of any microbe–host interactions.

Increased exoglycosidase activity has been detected in the gingival fluid of patients with periodontal disease (Beighton *et al.*, 1992). However, whether this increased enzyme activity is due to the action of microbial exoglycosidases, or is a result of the release of host enzymes as part of an inflammatory reaction, is unknown (Page, 1991). As a first stage in understanding the role of microbial exoglycosidases in gingival tissue degradation, and the perturbation of the host defence response, we have cloned the gene encoding the β -Nahase of *P. gingivalis* W83 in *Escherichia coli* and carried out a preliminary characterization of this enzyme.

METHODS

Bacterial strains, plasmids and media. Strains used in this study are listed in Table 1. *P. gingivalis*, *P. endodontalis* and *P. asaccharolytica* strains were routinely grown anaerobically in BM broth (Shah *et al.*, 1976), or on 7% (v/v) horse blood agar plates (Oxoid). *E. coli* strains were grown on L-agar or in L-broth supplemented where appropriate with 100 μ g ampicillin ml⁻¹ and 25 μ g tetracycline ml⁻¹. Plasmid pTTQ18 (Stark, 1987) was used for construction of the expression library of *P. gingivalis* W83 DNA.

DNA isolation and analysis. Chromosomal DNA was extracted from bacteria as described by Saito & Miura (1963). Large-scale plasmid purification was performed by the method of Clewell & Helinski (1969), while rapid small-scale purification was performed by the method of Birnboim & Doly (1979). Restriction endonucleases, calf-intestinal phosphatase (CIP) and T4 DNA ligase were obtained from BRL.

Nucleotide sequence analysis. Single-stranded M13 DNA templates were prepared from *E. coli* JM101 and sequenced by

the dideoxy chain-termination method of Sanger *et al.* (1977) using [α -³⁵S]thio-dATP and the modified T7 DNA polymerase, Sequenase version 2.0 (USB). The DNA fragments were analysed using buffer gradient gels (Biggins *et al.*, 1983). Nucleotide sequences were analysed using the Wisconsin (Devereux *et al.*, 1984), Lipman–Pearson (Lipman & Pearson, 1985) and Clustal-V (Higgins *et al.*, 1992) molecular biology programs on a VAX VMS cluster.

Construction and screening of a *P. gingivalis* expression library. *P. gingivalis* W83 chromosomal DNA was partially digested with *Sau*3A, and fragments from 1 to 4 kb purified on a sucrose-gradient (Milner *et al.*, 1993). Plasmid vector pTTQ18 (Stark, 1987) was linearized with *Bam*HI, treated with CIP, then ligated to the *Sau*3A-generated fragments, and the recombinant plasmids introduced into *E. coli* strain SURE by electroporation (Dower *et al.*, 1988). A representative *P. gingivalis* W83 genomic library was obtained by selecting ampicillin-resistant, recombinant white *E. coli* colonies on L-agar plates containing (0.004% w/v) X-Gal and 0.1 mM IPTG. β -Nahase-positive clones were detected as fluorescent colonies under long-wave UV light after overnight incubation on L-agar plates supplemented with 100 μ g 4-methylumbelliferyl-*N*-acetyl- β -D-glucosaminide (MUAG) ml⁻¹.

Southern blot analysis. Southern blots of restriction fragments were performed as described by Roberts *et al.* (1986). Radio-labelled probes were generated by extending hexadeoxynucleotide primers in the presence of [³²P]dCTP using the Klenow fragment of DNA polymerase I (Feinberg & Vogelstein, 1983). Hybridization was performed as described previously (Roberts *et al.*, 1986), and the filters then washed each for 15 min at 65 °C in 2 $\times$ SSC 0.1% SDS and then in 0.5 $\times$ SSC 0.1% SDS.

β -Nahase assays. Cultures (100 ml) of *E. coli* strain SURE harbouring the appropriate plasmid were incubated at 37 °C until an OD₆₀₀ of 0.5 was reached, at which point IPTG was added to a final concentration of 10 mM. The use of an increased concentration of IPTG (10 mM as opposed to 0.1 mM) was

Table 1. Bacterial strains

Strain	Relevant characteristics	Source
<i>E. coli</i>		
SURE	F' <i>proAB lacI^q lacZΔM15 Tn10</i>	Stratagene
JM101	F' <i>traD36 proAB lacI^q lacZΔM15</i>	I. S. Roberts
<i>P. gingivalis</i>		
W83	Clinical specimen	H. Shah*
WpH35	Clinical specimen	MRC Dental Unit, London
23A3	Clinical specimen	MRC Dental Unit, London
ATCC 33277	Type strain	ATCC
<i>P. endodontalis</i>		
ATCC 35406	Type strain	ATCC
<i>P. asaccharolytica</i>		
ATCC 8503	Type strain	ATCC

* Eastman Dental Hospital, Grays Inn Road, London.

required to maximize expression of the cloned genes (unpublished results). The cells were harvested at OD₆₀₀ of 1.0, resuspended in 3 ml 0.1 M MES buffer, pH 6.5, and sonicated four times for 15 s on ice with a 30 s cooling interval, using a Braun Labsonic 200 sonicator. *P. gingivalis* was grown in BM broth, resuspended in 0.1 M MES buffer, pH 6.5, and sonicated as described above for *E. coli*. Quantitative β -Nahase assays were performed in a final volume of 1 ml using a modification of the method of Casaregola *et al.* (1982). Briefly, this involved mixing 0.5 ml 0.01 M *p*-nitrophenyl-*N*-acetyl- β -D-glucosaminide, 0.3–0.48 ml 0.1 M MES buffer, pH 6.5, and 0.2–0.02 ml cellular sonicate on ice, followed by incubation at 37 °C for 1 h. The reaction was terminated by adding 3 ml 0.2 M sodium borate, pH 9.8, and the absorbance measured at 420 nm. Protein concentration was estimated using a protein estimation assay kit (Bio-Rad) with lysozyme as a standard. One enzyme unit was defined as the amount of enzyme which produced 1 mmol *p*-nitrophenol in 1 min, and activity is expressed as units (mg protein)⁻¹.

Radioactive labelling of proteins. Proteins were labelled *in vitro* with a prokaryotic DNA-directed transcription–translation kit (Amersham) in the presence of L-[³⁵S]methionine. The labelled samples were analysed using SDS-PAGE as described previously (Smith *et al.*, 1990).

RESULTS AND DISCUSSION

Cloning and expression in *E. coli* of a *P. gingivalis* gene encoding β -Nahase

The plasmid library of *P. gingivalis* W83 chromosomal DNA was screened for expression of β -Nahase activity on plates supplemented with MUAG. One positive (fluorescent) clone was identified from the library of 3500 clones. Plasmid DNA was extracted from the clone and used to re-transform *E. coli* strain SURE. All of the transformants expressed β -Nahase activity. This plasmid, termed pAL1, was then used for further studies. A restriction map of pAL1 was constructed (Fig. 1). The 4 kb *Sau3A* fragment cloned in pAL1 regenerated the *Bam*HI cleavage sites in pTTQ18, thereby allowing the entire insert to be removed on a *Bam*HI fragment (Fig. 1). Based on this restriction map, subclones and deletions were constructed utilizing the multiple cloning site of the vector. Plasmids pAL2 and pAL3 were constructed by sub-cloning a 2.7 kb *Xba*I and a 2.2 kb *Eco*RI fragment, respectively, into plasmid pTTQ18 (Fig. 1). Plasmid pAL4 was generated by deletion of the 2.2 kb *Eco*RI fragment. *E. coli* carrying pAL2 were positive for β -Nahase activity, whereas both pAL3 and pAL4 failed to express any appreciable β -Nahase activity (Table 2). This localized the gene encoding for β -Nahase activity to the 2.7 kb *Xba*I fragment (Fig. 1).

When pAL1 was introduced into *E. coli*, β -Nahase activity was detectable at levels lower than that with *P. gingivalis* W83 (Table 2). In *E. coli* harbouring pAL1, the expression of β -Nahase activity could not be increased by addition of IPTG (Table 2). This suggests that expression of β -Nahase activity in *E. coli* carrying pAL1 may be dependent on the *P. gingivalis* promoter. In contrast, addition of IPTG to *E. coli* carrying pAL2 induced a significant increase in β -Nahase activity (Table 2), indicating that in

pAL2 transcription of the gene encoding β -Nahase was now under control of the inducible *tac* promoter of pTTQ18.

Southern blot analysis

Southern blot analysis of suitably digested *P. gingivalis* genomic DNA was carried out using the 4 kb *Bam*HI fragment of *P. gingivalis* W83 chromosomal DNA (Fig. 1) as a radiolabelled probe. As predicted from the restriction map of pAL1 (Fig. 1), the probe hybridized to four *Pst*I fragments, two *Kpn*I fragments and two *Hind*III fragments of *P. gingivalis* W83 chromosomal DNA (Fig. 2). Chromosomal DNA from three other *P. gingivalis* strains digested with *Pst*I gave similar patterns of hybridization as *P. gingivalis* W83. The only restriction site polymorphism between the strains at this region of the chromosome was the presence of an additional fragment

Fig. 1. Restriction map of cloned *P. gingivalis* W83 chromosomal DNA. Thick solid lines indicate the DNA present in the various plasmids described in the text. The hatched area defines the *nahA* gene; N and C refer to the termini of the NahA protein. The circle labelled ptac denotes the *tac* promoter. Restriction sites are abbreviated as follows: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; K, *Kpn*I; P, *Pst*I; S, *Sph*I; X, *Xba*I.

Table 2. Expression of the *P. gingivalis nahA* gene

Strain	β -Nahase activity* [units (mg protein) ⁻¹]	
	– IPTG	+ IPTG
SURE(pTTQ18)	0.1 ± 0.01	0.2 ± 0.03
SURE(pAL1)	1.4 ± 0.1	1.5 ± 0.5
SURE(pAL2)	3.3 ± 1.2	23.3 ± 8.0
SURE(pAL3)	0.3 ± 0.2	0.4 ± 0.1
SURE(pAL4)	0.2 ± 0.1	0.3 ± 0.2
W83	22.5 ± 7.5	ND

* Mean ± SD of at least three experiments; ND, not done.

Fig. 2. Southern blot analysis using the radiolabelled *nahA* gene as a probe. Lanes 1, 2 and 3 contain *P. gingivalis* W83 chromosomal DNA digested by *Pst*I, *Hind*III and *Kpn*I, respectively. Lane 4 contains *P. asaccharolytica* chromosomal DNA and lane 5 contains *P. endodontalis* chromosomal DNA, both digested by *Pst*I. Lane 6 contains *P. gingivalis* WpH35 chromosomal DNA, lane 7 contains *P. gingivalis* 23A3 chromosomal DNA and lane 8 contains *P. gingivalis* ATCC 33277 chromosomal DNA, all digested by *Pst*I.

of approximately 0.3 kb in the chromosomal DNA of strain ATCC 33277 (Fig. 2). No sequence homologous to the probe could be detected in *Pst*I-digested genomic DNA of *P. asaccharolytica* ATCC 8503 or *P. endodontalis* ATCC 35406 (Fig. 2). Expression of β -Nahase activity is one of the discriminatory tests used to distinguish between *P. gingivalis* and the two other members of the genus *Porphyromonas* (Laughon *et al.*, 1982; van Winkelhoff *et al.*, 1985). The Southern blot results indicate that both *P. asaccharolytica* and *P. endodontalis* lack a gene homologous to the *P. gingivalis* structural gene for β -Nahase. Thus, the expression of β -Nahase activity would seem to be a reliable tool in the differentiation of *P. gingivalis* strains from other members of the genus.

Nucleotide sequence of the *nahA* gene and predicted amino acid sequence of NahA

Appropriate overlapping restriction fragments from the 2.9 kb *Kpn*I–*Bam*HI fragment of pAL1 (Fig. 1) were subcloned into M13mp18 and M13mp19, and the entire nucleotide sequence determined on both strands. Where necessary, oligonucleotide primers were synthesized and used to complete the nucleotide sequence on both DNA strands. The fragment was 2951 bp long with a G+C content of 47 mol%, which is comparable with the expected value for this species (46–48 mol%; Shah & Collins, 1988). It contained a single ORF of 2334 bp which encoded a 777-residue protein with a predicted molecular mass of 87 kDa (Fig. 3). We suggest that this gene be called *nahA* and the β -Nahase the NahA protein. Potential stem-loop structures were found both up- and

downstream of the *nahA* gene (Fig. 3). The potential stem-loop structure 5' to the *nahA* gene, which has an estimated ΔG of -24 kJ mol $^{-1}$, may act as a transcriptional terminator, and may explain why, in pAL1, expression of *nahA* was independent of the upstream *tac* promoter and could not be enhanced by addition of IPTG. Indeed, in pAL2, where *nahA* is cloned on an *Xba*I fragment (Fig. 1), this stem-loop structure is absent, and β -Nahase activity was increased by the addition of IPTG (Table 2). The second potential stem-loop structure, 3' to *nahA*, which has an estimated ΔG of -12.1 kJ mol $^{-1}$, is likely to take part in rho-independent transcriptional termination of *nahA* (Rosenberg & Court, 1979). No sequences homologous to the consensus sequence for an *E. coli* promoter were detected 5' to *nahA* (Fig. 3). This could explain the low levels of β -Nahase activity detected in *E. coli* harbouring pAL1 (Table 2). No obvious Shine–Dalgarno sequence was identified 5' to the putative start codon; however, this sequence is not essential for initiation of translation (McCarthy & Gualerzi, 1990).

The predicted amino acid sequence following the first ATG (Fig. 3) showed characteristic features of a bacterial lipoprotein signal sequence, having a positively charged amino-terminal segment followed by a hydrophobic sequence together with a processing site and a cysteine residue (Hayashi & Wu, 1990; Von Heijne, 1986). From the predicted amino acid sequence of the NahA protein, the site of action of the type II signal-peptidase would be the bond between the alanine at position 18 and the cysteine at position 19 (Fig. 3), with the mature form of NahA being acylated at this cysteine residue. It has been shown with both NlpA, an inner-membrane-associated lipoprotein of unknown function in *E. coli*, and the pullulanase enzyme of *Klebsiella oxytoca*, that an aspartic acid residue at position +2 (i.e. immediately after the acylated cysteine residue) acts as a lipoprotein sorting signal (Pugsley, 1993). The presence of an aspartic acid residue at this position directs a lipoprotein to the inner, rather than outer, membrane. Assuming such a sorting signal operates in *P. gingivalis*, the presence of a serine residue at this position in NahA (Fig. 3) may suggest that the protein is not localized in the inner membrane of *P. gingivalis*. The observation that β -Nahase activity is associated with the cell surface and outer membrane vesicles (OMV) released into the culture supernatant by *P. gingivalis* (Greenman & Minhas, 1989) would support the suggestion that the mature form of acylated NahA is associated with the outer membrane. OMV are highly proteolytic (Holt & Bramanti, 1991; Mayrand & Holt, 1988; Smalley & Birss, 1987) and it has been suggested that release of OMV may allow the targeting of hydrolytic enzymes to sites remote from the immediate environment of the cell (Holt & Bramanti, 1991). Obviously, the presence of β -Nahase within OMV would permit the degradation of GAG at remote sites and potentially allow the concerted action of both the β -Nahase and proteinase in the breakdown of proteoglycans and glycoproteins.

Protein database searches revealed homology between the predicted amino acid sequence of the NahA protein of *P. gingivalis* and a number of related hexosaminidases,

1	KpnI GGTACCGTTTCGATTCTCGCTCTGCATTCCAGACATCGGGCTTCCGCGTGGGTACTCCGGCCATCACCACTCGTGGCGTAAAAGAGATAAGATGGGCTAT	100
101	ATCTGGAGTTGATAGACCGTGTGCTCTCCGACCGGAGGACGAAGCCGTAATAGCATCGGTTCTGTACCGAAGTCAACCGGATGATGGCCGATTATCTCTC	200
201	TCTTTGCTTGGTAAGAAATCCCAAGAGTCTATCCCCCTATCCTTCTCTT XbaI AGGATAGGGGATAGTTCTTTACTGATATTTAGAAAGCACGCAAGGTTGT	300
301	GTGATGCCGGTGGTGCCCGGACACCTCACTTACGTGCTTTTTTGTGTTGATCACTAATTCATACTTGACAATGAAACGACTGACTTTCCGAACCATGCATT M K R L T F G A C I	400
401	TGCTGCCTCTGTCTCTTATGGCCTGCTCACAGAAAGCAAAGCAGGTGCAAAATCCCGAATACGACAAGGGTATAAACATCATTCCCTTCCGATGACGC C C L L S L M A C S Q K A K Q V Q I P E Y D K G I N I I P L P M Q L	500
501	TGACCGAATCGGACGACAGCTTTGAGGTGCGATGATAAGACCCTATCTCGGTATCTGCGCAAGAGCTAAAGCCTATCGCTAAACTTCTTCCGACAAAGCT T E S D D S F E V D D K T T I C V S A E E L K P I A K L L A D K L	600
601	AAGAGCATCAGCGACCTCTCTCCAGATAGAGATAGGCGAGGAGCCTTCGGGGAATGCTATTACATCGGTGTCGATACGGCTCTCTCTTAAAGAA R A S A D L S L Q I E I G E E P S G N A I Y I G V D T A L P L K E	700
701	GAGGGTTATATGCTCCGATCCGATAAGCGTGGTGTGCTATCATCGGCAAAATCTGCCCATGGTGTCTTCTACGGTATGACAGCTTTGCTCCAGCTCCTTC E G Y M S D K R G S I I G K S A H G A F Y G M Q T L L Q L L P	800
801	CTGCCGAAGTGAATCTTCGAATGAGGTACTGCTCCCCATGACGGTGCCCGCGTCGAGATCAAGGACGAACCGGCATTCCGGCTATCGTGCTTTATGCT A E V E S S N E V L L P M T V P G V E I K D E P A F G Y R G M L	900
901	GGATGTATGCCGTCATTTCCTTTCGGTGGAGGACATCAAGAAGCATATCGACATCATGGCCATGTTCAAGATCAATCGTTTCCATTGGCACCTGACAGAG D V C R H F L S V E D I K K H I D I M A M F K I N R P H W H L T E	1000
1001	GATCAGGCATGGCGTATCGAAATCAAGAAATACCCACGACTGACCGAAGTGGGGTCTACAAGGACGGAAGGGGACGGTACGCAGTACTCCGGTTTCTACA D Q A W R I E I K K Y P R L T E V G S T R T E G D G T Q Y S G F Y T	1100
1101	CGCAGGACCAAGTACGGGATATTGTACAATACGCATCGGATCATTTCATTCAGGTGATTCCCATGATCGAAATGCCCGGACATGCCATGGCTGCCCTCGC Q E Q V R D I V Q Y A S D H F I T V I P M I E M P G H A M A A L A	1200
1201	TGCTTATCCCGAGTTCCGTTGCTTCCACCGCGAATTCAGCCACGGATATTCTGGGGAGTGGAGCAGGATGTTTATTGTGCGCGTAAAGACAGCGTCTTC A Y P Q L E Q F R C F P R E F K P R I I W G V E Q D V Y C A G K D S V F	1300
1301	CGTTTTATCTCTGATGTTATCGACGAGGTAGCACCCCTTTTCCCGGCACATCTTCCATATCGGAGGGGACGAATGCCCTAAAGATCGATGGAAGGCTT R F I S D V I D E V A P L F P G T Y F H I G G D E C P K D R W K A C	1400
1401	GTTCGCTTTGTGAGAAGCGTATGCGTGACAATGGGTGAAGAGCAGAACGAGCTGCAGAGTTATTTCATCAACAAGCTGAAAAGGTCTTACAAAAGCA S L C Q K R M R D N G L K D E H E L Q S Y F I K Q A E K V L Q K H	1500
1501	CGGCAAGAGACTGATCGGTTGGGATGAAATCCTCGAAGCGGGGCTTGACCTTCTGCCACCGTTATGAGCTGGCGTGGAGGATGGTGGCATCGCAGCG G K R L I G W D E I L E G G L A P S A T V M S W R G E D G G I A A	1600
1601	GCTAATATGAATCAGATGTGATCATGACTCCGGGTAGCGGAGGTCTCTACTTGGATCATTTATCAGGAGATCCGACCGTCGAGCCTGTTGCCATCGGAG A N M N H D V I M T P G S G G L Y L D H Y Q G D P T V E P V A I G G	1700
1701	GTTATGCTCCATTGGAGCAAGTGTATGCTTACAATCCTTTGCCGAAAGAAATGCCCGGCCGATAAGCATCGCTACGTGCTCGGAGCACAGGCCAATCTGTG Y A P L E Q V Y A Y N P L P K E L P A D K H R Y V L G A Q A N L W	1800
1801	GGCAGAATACCTCTATACTTCCGAACGATACGACTATCAGGCTATCCAAAGCTACTGGCTGTGGCAGAGCTTACCTGGACACCGTGGCCAAAGAAAGAT A E Y L Y T S E R Y D Y Q A Y P R L L A V A E L T W T P L A K D	1900
1901	TTTCCGATTCTCTGTCGCCGTTTGGATAATGCCTGCTGCTGCTGACATGCATGGTATCAATTACCACATTCGCTGCCGGAACACCGGGTGGCTCTT F A D F C R R L D N A C V R L D M H G I N Y H I P L P E Q P G G S S	2000
2001	CCGACTTTATAGCCTTTACGGACAAGGCTAAGCTGACCTTCACGACATCGCGTCCGATGAAAAATGGTCTATACGCTGGACGAAACCGAACCATCCCTCAC D F I A F T D K A K L T F T T S R P M K M V Y T L D E T E P S L T	2100
2101	ATCGACTCCTTACACGGTCCCTCTGAATTTGCACAAACGGGCTCTCTGAAGATTCTGACCGTCACGGCCGCTGGGAAGATGAGTCCCGTACGCCGCATT S T P Y T V P L E F A Q T G L L K I R T V T A G G K M S P V R R I	2200
2201	CGTGTGGAGAAACACCTTCAATATGTCAATGGAAGTACCGGCACCGGAAACCGGACTGACCATTCGTACGGCTTACGGTACTTATATGATGTGCCTG R V E K Q P F N M S M E V P A P K P G L T I R T A Y G D L Y D V P D	2300
2301	ATCTGCAGCAGGTAGCCTCATGGGAAGTAGGACCGTTAGCTCTTTGGAGGAAATCATGCACGGGAAAGAGAAGATAACTTCTCCTGAAGTACTGGAGCG L Q Q V A S W E V G T V S S L E E I M H G K E K I T S P E V L E R	2400
2401	CAGAGTTGTAGAGCTACCGGTTATGTCTTATTCGGAGGATGGGTATATGAGTTCTCTACGGAAACACGAGTTTGGATTGATAATGTGAAGCTG R V V E A T G Y V L I P E D G V Y E F S T E N N E F W I D N V K L	2500
2501	ATCGACAATGTGGGCAAGTAAAGAAATCTCCCGTCGCAATAGCAGTCGTGCCCTTCAGAAAGCTACCATCCGATCAAGACGATATGGGTCCGAGCCA I D N V G E V K K F S R R N S S R A L Q K G Y H P I K T I W V G A I	2600
2601	TACAAGTGCCTGGCTACTTATTTGAACTACAGCAGGGAATGATACGGCTCAAGGGAGAAGAAAGTTCAAGCCGATCTCGTCCGATATGCTCTTTCA Q G A W P T Y W N Y S R V M I R L K G E E K F K P I S S D M L F Q	2700
2701	ATAAAGTCTCATCTTTATGGTTAAAGACGGCAACCTTTCCGACAGAAAAACGCAAGGCCCTAAATTAACCCCACTACCAGAAGAGGGTGTGTCAAAA	2800
2801	TTCAAGTTTAAAGAGGGGCGGGATCATCGTTTGTATAGAAAAACCGCTATCAATCAATAGTGTCTAAATGTTTGTCTTAGAGGGGGAGTTTGTAGTA	2900
2901	TGAGAGCTGAGTCGCTCCAATACAAGACTTTTGGGATTGAGGGGGATCC BamHI	2951

Fig. 3. Nucleotide sequence of cloned *P. gingivalis* W83 DNA. The nucleotide sequence is numbered 1 to 2951, with the predicted amino acid sequence of the NahA protein shown in single-letter code below the nucleotide sequence; restriction sites are shown in bold. Potential stem-loop structures are denoted by underlining of the nucleotide sequence. The predicted N-terminal signal sequence of the NahA protein is denoted by underlining the single-letter code.

* Homology determined using alignments generated as described by Lipman & Pearson (1985).

3404

protein visualized in the *in vitro* transcription-translation system is either a proteolytic cleavage fragment of NahA or may have arisen due to incorrect translation beginning at an internal ATG codon. The reason for the failure to obtain expression of any plasmid-encoded proteins in minicells harbouring either pAL1 or pAL2 is not clear. One possibility is that expression of *nahA* in minicells was lethal, since the presence of either pAL1 or pAL2 expressing β -Nahase activity in the *E. coli* minicell-producing strain, DS410, dramatically reduced the yield of minicells. Attempts to demonstrate that NahA is a processed lipoprotein by labelling *E. coli* cells harbouring pAL2 with [3 H]palmitic acid in the presence or absence of globomycin also proved unsuccessful (unpublished results).

Concluding remarks

The results reported here shed light on the nature of the β -Nahase of *P. gingivalis* and its probable location. This is the first step in the elucidation of the role of this enzyme in both the breakdown of host proteoglycans, and the removal of specific glycosidic residues from the surface of host cells and from immunologically important glycoproteins, such as IgG. We are currently studying the ability of recombinant NahA to degrade host proteoglycans. In addition, by generating β -Nahase⁻ mutants by site-directed gene replacement, and assaying the virulence of such mutants in appropriate animal models, it should be possible to elucidate the contribution of this enzyme in the pathogenesis of periodontal disease.

ACKNOWLEDGEMENTS

This work was supported by grants from the Medical Research Council of the UK. I. S. R. is a Lister Institute Research Fellow and gratefully acknowledges the support of the Lister Institute for Preventive Medicine.

REFERENCES

- Bapat, B., Ethier, M., Noete, K., Mahuran, D. & Gravel, R. A. (1987). Cloning and sequence analysis of a cDNA encoding the β -subunit of mouse β -hexosaminidase. *FEBS Lett* **237**, 191–195.
- Bartold, P. M. (1987). Proteoglycans of the periodontium: structure, role and function. *J Periodontal Res* **22**, 431–444.
- Bartold, P. M., Weibkin, O. W. & Thornard, J. C. (1982). Proteoglycans in human gingiva: molecular size distribution in epithelium and connective tissue. *Arch Oral Biol* **27**, 1–7.
- Beighton, D., Radford, J. R. & Taylor, M. N. (1992). Glycosidase activities in gingival crevicular fluid in subjects with adult periodontitis and gingivitis. *Arch Oral Biol* **37**, 43–48.
- Biggins, M. D., Gibson, T. J. & Hong, G. F. (1983). Buffer gradient gels and 35 S label as an aid to rapid DNA sequence determination. *Proc Natl Acad Sci USA* **80**, 3963–3965.
- Birnboim, H. C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **1**, 1513–1523.
- Boureau, H., Decre, D., Carlier, J. P., Guichet, C. & Bourlioux, P. (1993). Identification of a *Clostridium coelestium* strain involved in anti-*Clostridium difficile* barrier effect and determination of its mucin degrading enzymes. *Res Microbiol* **144**, 405–410.
- Bourgeau, G., Lapointe, H., Peloquin, P. & Mayrand, D. (1992). Cloning, expression and sequencing of a protease gene (*tpr*) from *Porphyromonas gingivalis* W83 in *Escherichia coli*. *Infect Immun* **60**, 3186–3192.
- Brown, C. A. & Mahuran, D. J. (1991). Active arginine residues in β -hexosaminidase. *J Biol Chem* **266**, 15855–15862.
- Cabezas, J. A. (1989). Some comments on the type references of official nomenclature (IUB) for β -N-acetylglucosaminidase, β -N-acetylhexosaminidase and β -N-acetylgalactosaminidase. *Biochem J* **261**, 1059–1061.
- Casaregola, S., D'Ari, R. & Huisman, O. (1982). Quantitative evaluation of *recA* gene expression in *E. coli*. *Mol & Gen Genet* **185**, 430–439.
- Clewell, D. B. & Helinski, D. R. (1969). Supercoiled circular DNA-protein complex in *Escherichia coli*: purification and conversion to an open circular form. *Proc Natl Acad Sci USA* **62**, 1159–1166.
- Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* **12**, 387–395.
- Dower, W. G., Miller, J. F. & Ragsdale, C. W. (1988). High efficiency transformation of *Escherichia coli* using electroporation. *Nucleic Acids Res* **16**, 6127–6145.
- Feinberg, A. P. & Vogelstein, B. (1983). A technique for radio-labelling restriction endonuclease fragments to a high specific activity. *Anal Biochem* **132**, 6–13.
- Graham, T. R., Zassenhaus, P. & Kaplan, A. (1988). Molecular cloning of the cDNA which encodes the β -N-acetylhexosaminidase from *Dictyostelium discoideum*. *J Biol Chem* **263**, 16823–16829.
- Greenman, J. & Minhas, T. (1989). Production of cell-bound and vesicle-associated trypsin-like protease, alkaline phosphatase and N-acetyl- β -glucosaminidase by *Bacteroides gingivalis* W50. *J Gen Microbiol* **135**, 557–564.
- Hayashi, S. & Wu, H. C. (1990). Lipoproteins in bacteria. *J Bioenerg Biomembr* **22**, 451–471.
- Higgins, D. G., Bleasby, A. J. & Fuchs, R. (1992). Clustal-V – improved software for multiple sequence alignment. *Comput Appl Biosci* **8**, 189–191.
- Holt, S. C. & Bramanti, T. E. (1991). Factors in virulence expression and their role in periodontal disease pathogenesis. *Crit Rev Oral Biol Med* **2**, 177–281.
- Jensen, H. K. & Ledet, T. (1986). Proteolysis of arterial basement membrane containing different amounts of carbohydrates. *Thromb Res* **44**, 47–53.
- Kato, T., Takahashi, N. & Kuramitsu, H. K. (1992). Sequence analysis and characterisation of the *Porphyromonas gingivalis* prtC gene, which expresses a novel collagenase activity. *J Bacteriol* **174**, 3889–3895.
- Koide, N., Matso, N. & Muramatso, T. (1977). Recognition of IgG by Fc receptor and complement: effects of glycosidase digestion. *Biochem Biophys Res Commun* **75**, 838–844.
- Korneluk, R. G., Mahuran, D. J., Neote, K., Klavinis, M. H., O'Dowd, F., Tropak, M., Willard, H. F., Anderson, M.-J., Lowden, J. A. & Gravel, R. A. (1986). Isolation of cDNA clones coding for the α -subunit of human β -hexosaminidase. *J Biol Chem* **261**, 8407–8413.
- Kytzia, H.-J. & Sandhoff, K. (1985). Evidence for two different active sites on human β -hexosaminidase A. *J Biol Chem* **260**, 7568–7572.
- Kytzia, H.-J., Hinrichs, U., Maire, I., Suzuki, K. & Sandhoff, K. (1983). Variant of GM₂ gangliosidosis with hexosaminidase A

- having a severely changed substrate specificity. *EMBO J* **2**, 1201–1205.
- Laughon, B. E., Syed, S. A. & Loesche, W. J. (1982). API ZYM system for identification of *Bacteroides* spp., *Capnocytophaga* spp., and spirochetes of oral origin. *J Clin Microbiol* **15**, 97–102.
- Li, J., Ellen, R. P., Hoover, C. I. & Felton, J. R. (1991). Association of proteases of *Porphyromonas (Bacteroides) gingivalis* with its adhesion to *Actinomyces viscosus*. *J Dent Res* **70**, 82–86.
- Lipman, D. J. & Pearson, W. R. (1985). Rapid and sensitive protein similarity searches. *Science* **227**, 1435–1441.
- Mayrand, D. & Holt, S. C. (1988). Biology of asaccharolytic black-pigmented *Bacteroides* spp. *Microbiol Rev* **52**, 149–153.
- McCarthy, J. E. G. & Gualerzi, C. (1990). Translational control of prokaryotic gene expression. *Trends Genet* **6**, 78–85.
- Milner, J. S., Dymock, D., Cooper, R. M. & Roberts, I. S. (1993). Penicillin binding proteins from *Erwinia amylovora*: mutants lacking PBP2 are avirulent. *J Bacteriol* **175**, 6082–6088.
- Nishikita, M., Yoshimura, F. & Nodasaka, Y. (1989). Possibility of *Bacteroides gingivalis* hemagglutinin possessing protease activity revealed by inhibition studies. *Oral Microbiol Immunol* **33**, 75–80.
- O'Dowd, B. F., Quan, F., Willard, H. F., Lamhonwah, A.-M., Korneluk, R. G., Lowden, J. A., Gravel, R. A. & Mahuran, D. J. (1985). Isolation of cDNA clones coding for the β -subunit of human β -hexosaminidase. *Proc Natl Acad Sci USA* **82**, 1184–1188.
- Page, R. C. (1991). The role of inflammatory mediators in the pathogenesis of periodontal disease. *J Periodontol Res* **26**, 230–242.
- Park, Y. & McBride, B. (1993). Characterisation of the *tpr* gene product and isolation of a specific protease-deficient mutant of *Porphyromonas gingivalis* W83. *Infect Immun* **61**, 4139–4146.
- Pugsley, A. P. (1993). The complete secretory pathway in Gram-negative bacteria. *Microbiol Rev* **57**, 50–108.
- Rafael, W. S. & Zyskind, J. W. (1989). The N,N' -diacetylchitobiase of *Vibrio harveyi*. *J Biol Chem* **264**, 14778–14783.
- Roberts, I. S., Mountford, R., High, N., Bitter-Suermann, D., Jann, K., Timmis, K. & Boulnois, G. J. (1986). Molecular cloning and analysis of genes for the production of the K5, K7, K12, and K92 capsular polysaccharides in *Escherichia coli*. *J Bacteriol* **168**, 1228–1233.
- Rosenberg, M. & Court, D. (1979). Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu Rev Genet* **13**, 319–354.
- Rouslahti, E. (1989). Proteoglycans in cell regulation. *J Biol Chem* **264**, 13369–13372.
- Saito, H. & Miuria, K. I. (1963). Preparation of transforming DNA by phenol treatment. *Biochim Biophys Acta* **72**, 619–629.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain termination inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.
- Shah, H. N. & Collins, M. D. (1988). Proposal for reclassification of *Bacteroides asaccharolyticus*, *Bacteroides gingivalis* and *Bacteroides endodontalis* in a new genus, *Porphyromonas*. *Int J Syst Bacteriol* **38**, 128–131.
- Shah, H. N. & Gharbia, S. H. (1989). Lysis of erythrocytes by the secreted cysteine proteinase of *Porphyromonas gingivalis* W83. *FEMS Microbiol Lett* **61**, 213–218.
- Shah, H. N., Williams, R. A. D., Bowden, G. H. & Hardie, J. M. (1976). Comparison of the biochemical properties of *Bacteroides melaninogenicus* from human dental plaque and other sites. *J Appl Bacteriol* **41**, 473–492.
- Slots, J. & Genco, R. J. (1984). Black-pigmented *Bacteroides* species, *Capnocytophaga* species and *Actinobacillus actinomycetemcomitans* in human periodontal disease: virulence factors in colonisation, survival and tissue destruction. *J Dent Res* **63**, 412–421.
- Slots, J., Bragd, L., Wikstrom, M. & Dahlen, G. (1986). The occurrence of *Actinobacillus actinomycetemcomitans*, *Bacteroides gingivalis* and *Bacteroides intermedius* in destructive periodontal diseases in adults. *J Clin Periodontol* **13**, 570–577.
- Smalley, J. W. & Birss, A. J. (1987). Trypsin-like enzyme activity of extracellular membrane vesicles of *Bacteroides gingivalis* W50. *J Gen Microbiol* **133**, 2883–2894.
- Smith, A. N., Boulnois, G. J. & Roberts, I. S. (1990). Molecular analysis of the *Escherichia coli* K5 *kps* locus: identification and characterisation of an inner membrane capsular polysaccharide transport system. *Mol Microbiol* **4**, 1863–1869.
- Stark, M. J. R. (1987). Multicopy expression vectors carrying the *lac* repressor gene for regulated high level expression of genes in *Escherichia coli*. *Gene* **51**, 255–267.
- Sundqvist, G., Carlsson, J., Herrmann, B. & Tarnvik, A. (1985). Degradation of human immunoglobulins G and M and complement factors C3 and C5 by black-pigmented *Bacteroides* species. *J Med Microbiol* **19**, 85–94.
- Takazoe, I., Nakamura, T. & Okuda, K. (1984). Colonisation of the subgingival region area by *Bacteroides gingivalis*. *J Dent Res* **63**, 422–426.
- Tipler, L. S. & Embery, G. (1985). Glycosaminoglycan-depolymerizing enzymes produced by anaerobic bacteria isolated from the human mouth. *Arch Oral Biol* **30**, 391–396.
- Uitto, V.-J. (1991). Extracellular matrix molecules and their receptors: an overview with special emphasis on periodontal tissues. *Crit Rev Oral Biol Med* **2**, 323–354.
- Von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res* **14**, 4683–4690.
- Watanabe, T., Oyanagi, W., Suzuki, K., Ohnishi, K. & Tanaka, H. (1992). Structure of the gene encoding chitinase D from *Bacillus circulans* WL-12 and possible homology of the enzyme to other prokaryotic chitinases and class III plant chitinases. *J Bacteriol* **174**, 408–414.
- Werries, E., Nebinger, P. & Franz, A. (1983). Degradation of biogene oligosaccharides by β -N-acetyl-hexosaminidase secreted by *Entamoeba histolytica*. *Mol Biochem Parasitol* **7**, 127–140.
- White, D. & Mayrand, D. (1981). Association of oral *Bacteroides* with gingivitis and adult periodontitis. *J Periodontol Res* **16**, 259–265.
- van Winkelhoff, A. J., van Steenberghe, T. J. M., Kippuw, N. & De Graaf, J. (1985). Further characterisation of *Bacteroides endodontalis*, an asaccharolytic black-pigmented *Bacteroides* species from oral cavity. *J Clin Microbiol* **22**, 75–79.

Received 20 June 1994; revised 1 August 1994; accepted 15 August 1994.