GINGIPAINS FROM *PORPHYROMONAS GINGIVALIS* – COMPLEX DOMAIN STRUCTURES CONFER DIVERSE FUNCTIONS

N. Li and C. A. Collyer*

School of Molecular Bioscience, University of Sydney, NSW, Australia

Gingipains, a group of arginine or lysine specific cysteine proteinases (also known as RgpA, RgpB and Kgp), have been recognized as major virulence factors in *Porphyromonas gingivalis*. This bacterium is one of a handful of pathogens that cause chronic periodontitis. Gingipains are involved in adherence to and colonization of epithelial cells, haemagglutination and haemolysis of erythrocytes, disruption and manipulation of the inflammatory response, and the degradation of host proteins and tissues. RgpA and Kgp are multi-domain proteins composed of catalytic domains and haemagglutinin/adhesin (HA) regions. The structure of the HA regions have previously been defined by a gingipain domain structure hypothesis which is a set of putative domain boundaries derived from the sequences of fragments of these proteins extracted from the cell surface. However, multiple sequence alignments and hidden Markov models predict an alternative domain architecture for the HA regions of gingipains. In this alternate model, two or three repeats of the so-called “cleaved adhesin” domains (and one other undefined domain in some strains) are the modules which constitute the substructure of the HA regions. Recombinant forms of these putative cleaved adhesin domains are indeed stable folded protein modules and recently determined crystal structures support the hypothesis of a modular organisation of the HA region. Based on the observed K2 and K3 structures as well as multiple sequence alignments, it is proposed that all the cleaved adhesin domains in gingipains will share the same β -sandwich jelly roll fold. The new domain model of the structure for gingipains and the haemagglutinin (HagA) proteins of *P. gingivalis* will guide future functional studies of these virulence factors.

Keywords: *Porphyromonas gingivalis*, cleaved adhesin, gingipain, haemagglutinin, lysine- and arginine-specific cysteine proteases

Introduction

Periodontal disease

Periodontal diseases are caused by bacterial infections in the periodontium which is composed of gingiva, periodontal ligament, cementum and alveolar bone, all of which support and anchor the teeth [1]. Among these diseases, gingivitis and chronic periodontitis are the two most common forms of periodontal diseases. Gingivitis is non-destructive to the tooth tissues and most people can recover from it through effective oral hygiene which removes plaque. However, gingivitis can further progress to destructive periodontitis in some patients perhaps in part due to their unique host responses to the pathogenic organisms in the plaque and other factors [2]. Chronic periodontitis is the most common form of periodontitis [3]. It has been defined as “an infectious disease resulting in inflammation within the supporting tissue of the teeth, progressive de-attachment and bone loss” [4]. The inflammation of the gums has been considered to be a result of a direct host immune response to the increasing amount of bacterial pathogens in the dental plaque formed on the gingival margin [2, 5].

Recently, public health concerns have been raised as more and more evidence accumulates to demonstrate that periodontal diseases are also associated with cardiovascu-

lar diseases [6], diabetes mellitus [7–8], rheumatoid arthritis [9–10] and Alzheimer disease [11–13].

The generally accepted specific plaque hypothesis considers that only a selected few species among several hundred oral microorganisms are actively pathogenic and they are responsible for the inflammatory events and the destructive processes [14–16]. These studies have shown that only 10–20 bacterial pathogens in the dental plaque, the so-called periodontopathic microbiota, are responsible for destructive periodontitis, including *Porphyromonas gingivalis*, *Bacteroids forsythus*, *Treponema denticola*, *Prevotella intermedia*, *Prevotella nigrescens*, *Fusobacterium nucleatum*, *Peptostreptococcus micros*, *Campylobacter rectus*, *Treponema pectinovorum*, *Treponema vincentii*, *Eikenella corrodens* and *Actinobacillus actinomycetemcomitans* [15, 17–19]. Among these pathogens, *Porphyromonas gingivalis* has been considered to be a major contributor to the development of chronic periodontitis [20–22] and has been widely studied.

An unusual oral pathogen, *Porphyromonas gingivalis*

P. gingivalis is commonly detected in chronically inflamed periodontal lesions and the proportion of this bacterium in the anaerobically cultivable flora in subgingiva can be as

*Corresponding author: Charles A. Collyer; Phone: +61 293512794; Fax: +61 293514726;
E-mail: charles.collyer@sydney.edu.au

high as 50% [23]. *P. gingivalis* is a non-motile Gram-negative anaerobic bacillus [18]. In contrast to other members of the genus *Porphyromonas*, many of which are able to grow on complex carbohydrates, *P. gingivalis* is asaccharolytic and reliant upon nitrogenous substrates such as proteins or peptides as nutrients and for metabolic energy [24–26]. *P. gingivalis* is able to produce a large amount of proteinases to degrade proteins from host or other microorganisms in order to meet its special nutritional requirements [5, 18]. Importantly, this bacterium needs exogenous haem for growth due to the lack of a haem biosynthesis pathway [27–28]. *P. gingivalis* is mostly found in bleeding chronic periodontal lesions, where haemoglobin from ruptured erythrocytes provides a very convenient and abundant haem source. When growing on blood agar plates *P. gingivalis* colonies are initially white to creamy in colour, but turn dark red to black after 6–10 days [26]. The black pigmentation has been verified as an accumulation of iron (III) protoporphyrin IX in the form of the μ -oxo dimer $[\text{Fe(III)PPIX}]_2\text{O}$ on the bacterial cell surface [29]. Among the common laboratory strains and clinical isolates of *P. gingivalis*, strains W83, W12 and W50 are found to be more virulent than strains 381, HG66 and ATCC33277 [30].

The major virulence factors of *P. gingivalis* include fimbriae, capsule, outer membrane vesicles, lipopolysaccharide (LPS), toxic metabolites and proteinases [18, 24–25].

P. gingivalis expresses a group of endopeptidases called gingipains on the outer membranes which are responsible for at least 85% of the proteolytic activity and 99% of the “trypsin-like” activity produced by the bacterium [25]. The primary aim of the expression of these proteinases by this pathogen is to digest proteins for nutrition, but gingipains are also found to be involved in the destruction of the host periodontal matrix and alveolar bone, host cell adhesion and invasion, and in dysregulation of the host immune response [5].

The gingipain proteins

Gingipains are cysteine proteinases that belong to the peptidase family C25 [31]. There are three types of gingipains in *P. gingivalis*: lysine-specific gingipain (Kgp), arginine-specific gingipain A (RgpA), and arginine-specific gingipain B (RgpB) [32]. Gingipains are principally located on the outer membranes and outer membrane vesicles of all *P. gingivalis* strains except for HG66, which produces and secretes soluble forms of gingipains into the extracellular milieu [5, 25, 33–34]. Genomic studies show that gingipains are encoded by individual gene loci (*kgpA*, *rgpA* and *rgpB*), found in the genomes of all *P. gingivalis* strains [35–39]. The proteins encoded by *kgp* or *rgpA* genes consist of a signal peptide, an N-terminal pro-fragment, a Lys-specific or Arg-specific catalytic domain and a large C-terminal haemagglutinin/adhesin (HA) region (Fig. 1) [35, 40]. In contrast, the protein encoded by the *rgpB* gene consists of a signal peptide, an N-terminal pro-fragment, an Arg-specific catalytic domain and a small C-terminal fragment [41]. The C-terminal fragment of the RgpB protein shows a significant deletion in the HA region when compared with the C-terminal region in the RgpA protein [41]. In addition to the catalytic domains the differences between these proteins are found in the HA regions. It is thought that the adhesion properties of these varying regions confer the biological specificities of these proteins.

The gingipain domain structure hypothesis

Sequence analyses of purified gingipains

Studies of the possible domain organization of the structures of gingipains have been conducted by analysing the proteins purified from culture media and outer membranes

Fig. 1. Schematic diagram of the gingipain domain structure hypothesis. The colour matches indicate that the sequences are highly similar. This figure was adapted from [38, 42, 44–45]

using SDS-PAGE, peptide mass fingerprinting and N-terminal sequencing [42–44]. The proposed gingipain domain structure hypothesis which is supported by genetic analysis of cloned *kgp* and *rgp* from strain W50 as well as from the translated peptide sequences [38, 43] is presented in *Figure 1*. These peptides were also designated as HA1, HA2, HA3 and HA4 by DeCarlo et al. in their studies (*Fig. 1*) [45]. Further details on the putative domain boundaries have been provided by subsequent analysis of the outer membrane proteins from *P. gingivalis* W50 using the combined techniques of two dimensional gel electrophoresis, N-terminal sequencing and peptide mass fingerprinting [44]. It was proposed that the Kgp44 region might consist of three parts: Kgp14, Kgp13 and Kgp20 (*Fig. 1*) [44]. Moreover, it was observed that gingipains are present as

noncovalently associated complexes because purified gingipains only fragmented into to smaller masses when they were boiled [42]. The presence of various detergents alone was insufficient to mediate this separation [46].

Sequence homologies of gingipains from different *P. gingivalis* strains

The corresponding translated amino acid sequences of RgpA or RgpB proteins found in different strains are more than 98% identical [47–48]. The catalytic domains of RgpA (RgpA_{cat}) and RgpB (RgpB_{cat}) proteins typically share ~90% sequence identity, but the pro-peptides are only ~75% identical in these two proteins (*Fig. 2*).

Fig. 2. Schematic diagram summarizing the gingipain domain structure hypothesis for gingipains and HagA from the HG66, W50, 381, W83 and W83v strains of *P. gingivalis*. The proteins are divided into domains of sequence similarity bounded by proteolytic processing sites and coloured and labelled accordingly. Arrows indicate the proteolytic processing sites. The sequences in the regions between two vertical or slant dash lines are compared and the amino acid sequence identity in percentage is indicated. Similar sequences are indicated by the same colour. A part of the data was adapted from [25, 53]

In comparison with other gingipains, the catalytic domain of Kgp (Kgp_{cat}) is only 20–30% identical to those of RgpA_{cat} and RgpB_{cat}. However, in the C-terminal regions of both RgpA_{cat} and Kgp_{cat}, there is a highly conserved sequence containing 31 residues. This conserved sequence contains the adhesin binding motif 1 (ABM1), which is also in part observed in the Rgp44/Kgp39 regions and the Rgp17/Kgp44 regions [38, 49].

Despite the differences in the catalytic domains, RgpA from strain HG66 and Kgp from strain HG66 have almost identical sequences in the HA regions with the exception of the N-terminal 146 residues in the Kgp39 and Rgp44 fragments, respectively. The Kgp proteins from strains W12, W50 and W83 are 99% identical in sequence, but differ from the Kgps in strains HG66 and 381 in the Kgp44 regions [5] (Fig. 2). The catalytic domain of Kgp W83v lacks the pro-peptide and part of the N-terminal sequence, while the remaining sequence is identical to that of KgpW83 except for a 154-residue fragment in the Kgp39 domain (Fig. 2). Strains HG66, 381, W83 and W83v are considered to be distinct lysine gingipain (*kgp*) biovars due to their different *kgp* gene structures [50].

The Kgp15 and Rgp15 regions in all of the gingipains are virtually identical (sharing 97% sequence identity) with only a few residue substitutions including G1186/D, S1204/T, S1208/A and N1288/D as numbered when found in Kgp W83. The last 72 residues in the C-terminus of RgpB are found to be partially conserved in the C-termini of RgpA and Kgp. This conserved sequence has been suggested to function as an anchor that attaches the gingipains to the cell outer membrane. The soluble form of RgpB does not contain this peptide and it is probably removed by proteolysis [31, 34]. This 72-residue peptide has also been suggested to play critical roles in maintaining the correct folding of the catalytic domain of RgpB and the posttranslational glycosylation of this protein [51].

The C-terminal HA sequences have also been found in the haemagglutinin, HagA (Fig. 2). Han et al. [52] have reported that the *hagA* gene of *P. gingivalis* strain 381 encodes four large contiguous direct peptide repeats: Y416-D868, Y869-D1324, Y1325-D1780 and Y1781-D2236. These four peptide repeats share 94–100% sequence identity and are 95% identical to the fragment Y534-D992 found in Kgp W83v. In addition, the region of Y2237-K2628 in the C-terminus of HagA is 97% identical to the region of Y1341-K1732 in the C-terminus of Kgp W83 (Fig. 2).

The structure of the catalytic domain of RgpB

The structure of only one catalytic domain in the gingipain protein family, the crystal structure of RgpB, has been reported. The RgpB protein was purified from the culture medium of *P. gingivalis* strain HG66 in a soluble form and the crystal structure represents an archetypical gingipain catalytic domain consisting of 435 residues [31].

The overall structure of RgpB was described as a “crooked one-root tooth” with a spherical “crown” as the

N-terminal catalytic domain (351 residues) and the “root” being the C-terminal Ig-like domain (84 residues) (Fig. 3). The catalytic domain is composed of two sub-domains A and B with similar α/β open sheet topologies formed by a central β -sheet and two α -helices flanked on either sides. The structure of RgpB in complex with D-Phe-Phe-Arg-chloromethylketone (FFRCMK) inhibitor (PDB entry 1CVR) has also been reported [31]. The position of the FFRCMK inhibitor defines an active site formed by H211 and C244 in sub-domain B. The Cys-His catalytic dyad acts to bind and cleave the Arg-Xaa substrate. The residue Arg in FFRCMK forms a covalent bond with C244 Sγ via its methylene group [31]. The side chain carboxylate group of residue D163 in RgpB forms a salt bridge with the guanidyl group of the Arg of the FFRCMK inhibitor. Several hydrogen bonds connect the inhibitor to residues G212, Q282 and W284 which are also close to the active site. A Zn²⁺ ion is observed close to the active site interacting with H211 and E152 but only in the RgpB-FFRCMK complex structure. Although E152 in RgpB is replaced by an Asp in Kgp, G212 and the catalytic residues H211 and C244 are conserved in the Kgp catalytic domain sequence [31]. This suggests that similar active sites might be shared by these two enzymes despite the fact that they share only about 22% overall se-

Fig. 3. Representation of the structure of RgpB in complex with D-Phe-Phe-Arg-chloromethylketone (FFRCMK) inhibitor (PDB entry 1CVR). The RgpB and FFRCMK [31] are represented as ribbon and sticks, respectively. The sub-domains A and B of the catalytic domain are coloured in blue and deep-teal, respectively, the C-terminal Ig-like domain is in red colour, and the FFRCMK is coloured orange. Six Ca²⁺ and two Zn²⁺ ions are shown as spheres and coloured gold and green, respectively. The side chains of residues E152, H211 and C244 are shown as red sticks

quence identity [5]. The C-terminal domain of RgpB has a β -barrel topology formed by seven β -strands which is similar to the members of the immunoglobulin superfamily (IgSF). IgSF domains are known to be involved in the recognition, binding and adhesion processes of cells [31].

The closest structural homologues of RgpB are the caspases (Asp-specific cysteine proteases), in particular caspase-1 and caspase-3, with their catalytic domains being partially superimposable. Residues H211 and C244 found in the active site of RgpB are structurally conserved in the active sites of these caspases, indicating that these enzymes might be evolutionally related [31]. The catalytic domains of gingipains are chemically highly specific and cleave polypeptides at either lysine or arginine residues while the biologically relevant targets are selected *in situ* by the adhesion/binding properties of the associated HA regions.

Investigating the structure of the HA region of gingipains

Until recently attempts using recombinant proteins to understand the structure and function of the HA regions have focussed upon the gingipain domain structure hypothesis. For example, the recombinant Kgp15/Rgp15 protein construct has been studied extensively. Proposals have been made regarding direct haemoglobin and haem/haemin interactions being mediated by this polypeptide [29, 54–56]. However, structural studies of this protein construct have been unsuccessful, possibly because it tends to aggregate when it is concentrated above 1 mg/ml in solution (D. Langley, personal communication). Also, recombinant Rgp17 and Rgp27 of RgpA from *P. gingivalis* ATC33277 have been purified in very low yields (30–40 μ g purified protein per litre culture) and the recombinant product appears not to be suitable for structural studies [57].

Despite these studies of the binding activities of synthetic protein constructs, most being designed as fragments based upon the gingipain domain structure hypothesis, there is presently no clear understanding of the contribution made by the HA regions to gingipain function. The complexity of the HA structure and the consequent practical challenges of its protein biochemistry is a barrier that has until recently impeded this research effort.

The biological functions of gingipains

Gingipains in adhesion and haemagglutination

P. gingivalis is able to agglutinate erythrocytes and this is considered to be evidence that *P. gingivalis* adheres to host tissue [58–61]. Fimbriae and LPS have both been implicated as factors in the adherence of *P. gingivalis* to epithelia, erythrocytes and other bacteria [62–65]. These factors are complemented by the outer membrane expressed gingipains and HagA proteins which also play important roles in cell adhesion.

It has been observed that both intact *P. gingivalis* and purified gingipains from cell membranes or growth culture

are able to agglutinate erythrocytes and bind to haemoglobin [46]. The observation that two triple mutants, *rgpA⁻rgpB⁻kgp⁻* and *rgpA⁻kgp⁻hagA⁻*, lack haemagglutination and haemoglobin binding ability indicates the importance of the HA regions in mediating these activities [66].

Extensive studies have been carried out to identify the colonization and haemagglutination motifs within the gingipain amino acid sequences. Monoclonal antibody MAb 61BG1.3 raised against formalinized *P. gingivalis* W83 cells, was shown to be able to inhibit host re-colonization in immunized periodontal patients for 6–9 months [67–69]. Using serially truncated recombinant proteins and western blots, the “colonization epitope” recognized by MAb 61BG1.3 was mapped to residues G907–T931 of RgpA [69–70]. The peptide G907–T931 was recognised as a haemagglutination motif by using truncated recombinant proteins and another monoclonal antibody MAb 1A1, raised against formalinized *P. gingivalis* DCR 2015, inhibit the haemagglutination mediated by *P. gingivalis* W50 culture supernatant [68]. Later, a small peptide “PVQNLT”, which presents within the G907–T931 sequence was characterized as a haemagglutinin-associated short motif by using monoclonal antibody MAb-Pg-vc raised against outer membrane vesicles of *P. gingivalis* strain 381 [71]. This short motif is present as “PVQNLT” or “PVKNLK” at the C-terminus end of RgpA_{cat} and Kgp_{cat} domains, in the Rgp44 and Kgp39 regions and in the Rgp17 and Kgp44 regions. Another synthetic peptide containing residues G1083–T1102 of RgpA strongly inhibited the haemagglutination mediated by intact W50 cells with a minimum inhibitory concentration of <1.5 μ M [70]. Moreover, synthetic peptides spanning residues P785–K840, A934–N973, G974–F1042 and R1093–V1112 of RgpA are all recognized by antibodies within human sera taken from patients with periodontitis [70].

Recently, Sakai et al. found that a recombinant full length Rgp44_{720–1138} protein (referred to as rHGP44A) did not display haemagglutination activity [59] even though it contained a previously identified haemagglutinating determinant (residues G1083–T1102) [70]. Surprisingly, a second truncated form, Rgp44_{720–1081} (referred to as rHGP44B), did, however, display haemagglutination activity. A synthetic peptide corresponding to Rgp residues G1083–T1102 bound to Rgp44_{720–1081} in a dose-dependent manner and subsequently inhibited the haemagglutination mediated by Rgp44_{720–1081} [59]. It was suggested that the peptide G1083–T1102 functioned as a haemagglutination suppressor rather than a haemagglutination determinant [59]. Additionally, an anti-Rgp44_{720–1138} antibody was found to block the haemagglutination mediated by *P. gingivalis* with a range of concentrations of 2.5–10 μ g/well [59]. In contrast, anti-HbR, an antibody raised against the Rgp15/Kgp15 region, did not block *P. gingivalis*-mediated haemagglutination, suggesting that Rgp44 region played more important role in haemagglutination than the Rgp15/Kgp15 region [59].

In addition to ABM1 which was proposed to mediate the formation of RgpA–Kgp complexes [38], O’Brien-Simpson et al. have characterised another two adhesin binding motifs, ABM2 (S886–A914 in Kgp) and ABM3

(V939-F972 in RgpA), using competitive ELISA and antibodies raised against synthetic peptides [49]. The anti-ABM2 antibody inhibited the binding of RgpA-Kgp complexes to fibrinogen, fibronectin and collagen type V but did not inhibit the binding to haemoglobin. In contrast, the anti-ABM3 antibody only inhibited the binding of gingipain complexes to haemoglobin but not to the other proteins [49]. The epitopes in ABM2 and ABM3 recognized by these antibodies were further identified to be the short motifs "EGLTAT-TFEEDGVAA" and "GTPNPNPNPNPNPGT", respectively [49]. The binding of gingipains to fibrinogen, fibronectin and laminin were also reported by Pike et al. [72], and the apparent disassociation constants (K_D) of Kgp and RgpA binding to fibrinogen were estimated to be 4 nM and 8.5 nM, respectively [72]. The sequence "EGLTAT-TFEEDGVAA" presents three times in both the Kgp and RgpA sequences [49]. The "FEED" motif has been reported to be an essential sequence for proteins binding to fibronectin [73]. These colonization and haemagglutination determinants, and ABM1, ABM2 and ABM3 peptides are also found in HagA sequences [49, 53].

Gingipains in haemolysis

The haemagglutination mediated by *P. gingivalis* is a primary step for this pathogen in its adherence to erythrocytes. *P. gingivalis* is known to lyse erythrocytes and the key haemolysin has been identified as being associated with the outer membrane gingipains [74]. Both Rgp and Kgp have been reported to possess haemolytic activity [61, 75], but the receptors on the erythrocytes are still not known. The band 3 protein from the membranes of erythrocytes has been considered to be one possible target because the purified band 3 protein can bind to *P. gingivalis* cells and inhibits the haemagglutination mediated by a protein purified from the culture supernatant of strain W83 [76]. Sakai et al. reported that a recombinant protein Rgp44₇₂₀₋₁₀₈₁ was able to bind to the erythrocyte membrane protein glycophorin A with a K_D of 30 μ M [59].

Gingipains in haem acquisition

P. gingivalis has an obligate requirement for exogenous iron and protoporphyrin IX (PPIX) for its growth due to gene defects. Genes *hemA*, *hemL*, *hemB* and *hemC* encoding the enzymes that are involved in haem biosynthesis pathway cannot be identified in the genome of *P. gingivalis* [27]. While *P. gingivalis* can utilize haem from other sources such as serum albumin, transferrin, lactoferrin, haemopexin, catalase, lactoperoxidase and cytochrome [77–78], haemoglobin is the preferred haem and iron source [79].

P. gingivalis is known to be able to accumulate haem on the cell surface and the black pigmentation is a typical phenotype for this pathogen when grown on blood agar plates [80–81]. The *kgp*[−] mutants lack the black pigmentation, indicating that Kgp is involved in haem capture [66,

75]. Moreover, Lewis et al. [75] have confirmed that the *kgp*⁺ *rgpA*[−] mutant strain has higher hydrolysis efficiency of haemoglobin than the *kgp*[−] *rgpA*⁺ mutant strain. Secondly, purified Kgp from W83 vesicles have a haemaglobinase activity which is comparable to haemoglobin hydrolysis by intact outer membrane vesicles which contain both active Kgp and Rgp [75]. Therefore, Kgp has been designated as the major active haemaglobinase in *P. gingivalis* [75].

The haemoglobin binding abilities of RgpA and Kgp have been examined using both purified monomeric and heterodimeric proteins [44]. Both forms of RgpA and Kgp bind strongly to haemoglobin with dissociation constants (K_D values) of ~2–4 nM [23]. This binding affinity can be destroyed by boiling, indicating that the haemoglobin binding activity of gingipains requires folded proteins [45]. Both purified RgpA and Kgp have been observed to bind to haemin [45, 82]. In contrast, RgpB, which lacks the HA region, has little or no binding activity to either haemin or haemoglobin, further confirming that the HA regions play a critical role for gingipains to bind with haemoglobin and haem [82].

The haemoglobin-binding receptor in gingipains has been identified as Kgp15/Rgp15 region [54, 83], which is also referred to as the HA2 domain [45]. The Kgp15/Rgp15 region consists of 135 residues with a theoretical molecular weight of 15 kD, but the recombinant protein (referred to as rKgp15 or rRgp15) runs anomalously as a single band at 19 kD on the SDS-PAGE under reducing conditions [54–55]. rKgp15 has been found to bind to haemoglobin in a pH-dependent manner with a K_D of 50 nM at pH 5.5 [54], and the rRgp15 binds to haemoglobin and haemin with K_D values of 4 nM and 16 nM, respectively [45].

Studies by Smalley et al. revealed that rKgp15 is able to convert monomeric haem (Fe(III) PPIX.OH) to μ -oxo bishaem ([Fe(III)PPIX]₂O), the main component in the black pigment deposited on the cell surface by *P. gingivalis* [29]. The accumulation of μ -oxo bishaem on the cell surface is proposed to be a defence mechanism used by *P. gingivalis* against the ingress of oxygen and reactive oxygen species such as H₂O₂, and a mechanism for iron and haem storage [84].

Olczak et al. have examined the binding of gingipains with haemoglobin, non-metal porphyrins and metal porphyrins [82]. They found that Kgp and RgpA bound more tightly to haemoglobin than to the porphyrins, implicating the involvement of different mechanisms in regulating the binding capacity of gingipains to haemoglobin and porphyrin. It was suggested that the gingipain–haemoglobin binding might occur via protein–protein interactions [82]. In contrast, DeCarlo et al. have observed that the binding of rKgp15, RgpA and Kgp to haemoglobin and haemin was inhibited by porphyrin PPIX [45]. Further studies using modified porphyrins suggested that rKgp15 and porphyrin recognition is mediated via the vinyl side chain [55].

Accumulated evidence supports a critical role for Kgp in binding of and in the degradation of haemoglobin, haem capture and the accumulation of the black pigmentation [54, 75, 82]. However, recent studies by Smalley et al. in-

licated a specific role for RgpA in the black pigmentation of *P. gingivalis*, as that the pigmentation process requires both Rgp and Kgp activities [85]. They have proposed that Kgp and Rgp co-operate to degrade haemoglobin, with where RgpA mediating the conversion of oxyhaemoglobin to methaemoglobin and Kgp rapidly degrading methaemoglobin to generate μ -oxo bisheam [86].

Several mechanisms have been proposed to explain haem transportation in *P. gingivalis*. Olczak et al. have demonstrated that gingipains are able to bind to haem receptor proteins such as HmuR [82]. They have proposed that haemoglobin is degraded by Kgp, upon which the released haem is then initially captured by gingipains and delivered to the outer membrane receptor HmuR [82]. Recently, Gao et al. have identified another haemophore-like haem-binding protein, HusA, which can bind to dimeric haem with a K_D of ~ 1 nM [87]. These authors have proposed that the haem transportation involves the following steps: 1) Kgp degradation of host methaemoglobin and capture of the released monomeric haem; 2) Kgp converts the monomeric haem to μ -oxo dimeric haem; 3) the dimeric haem is transferred to the haem receptor HusA and subsequently transferred to transmembrane protein HusB; 4) the dimeric haem is transported to the periplasm by HusB [87]. Although these proposals provide some biochemical and molecular insight into haem transportation, the detailed mechanisms need yet to be further elucidated.

Gingipains in the degradation of host proteins and in house keeping

Gingipains, the “trypsin-like” cysteine proteinases, are capable of cleaving or degrading a range of host proteins including plasma proteins, extracellular matrix proteins, cytokines and host cell surface proteins [25]. Fifteen *P. gingivalis* strains grow well in KGB minimal medium containing bovine serum albumin as the sole carbon and nitrogen source, and W50 and ATC33277 also grow well in a medium with bovine γ -immunoglobulin as the sole carbon and nitrogen source [88]. However, strains with *rgpA* or *kgp* gene deficiencies and an *rgpA*⁻*kgp*⁻*RgpB*⁻ triple mutant do not grow in media supplied with human serum as the sole carbon and energy source, indicating the important roles of RgpA and Kgp in degradation of human serum proteins such as albumin, transferrin and IgG [89]. The binding and degradation of haemoglobin by gingipains is similarly linked to acquisition of peptide nutrients and the essential haem.

Aside from the dual roles of nutrient release and haem capture, gingipains are also responsible for the housekeeping work of gingipain maturation. Almost all of the proteolytic processing sites in gingipains and HagA are Arg-Xaa bonds except for one Lys-Xaa site that is found at the C-terminus of Kgp15/Rgp15. Perhaps this indicates the greater importance of RgpA and RgpB than Kgp in maturation and self-processing [25]. Gingipains also process profimbrilin to fimbrilin, the building block subunit used in the assembly of the fimbria [90].

Gingipains and the dysregulation of the host immune response

As an intruder to the human body, *P. gingivalis* is naturally a target of the host immune system including the initial non-specific innate immune response and the specific adaptive immune response [5]. The innate immune system provides the first-line of defence against infections by activating the complement system and employing antibacterial peptides, neutrophils and tissue-resident macrophages. The adaptive immune system, which is activated by the innate immune response subsequently defends the host against bacterial invasion by using antibodies, activated macrophages and cytotoxic T cells [34]. To survive a vigorous attack from the host immune system, *P. gingivalis* manipulates both the layers of host defense and gingipains are key players in the process.

Cationic antimicrobial peptides including human β -defensins, cathelicidin LL-37 peptide and neutrophil-derived α -defensins form the first line of innate immune defence in the oral cavity. Pathogen invasion induces the expression of these antimicrobial peptides [91]. On the other hand, gingipains have been found to degrade these induced antimicrobial peptides, enabling the survival of this pathogen in the periodontal lesions [34].

The innate immune responses summon leukocytes, mainly neutrophils and tissue-resident macrophages, to eliminate the intruding pathogens. Gingipains have been found to mediate a dysfunction in the phagocytes [5] by cleaving the formyl-methionyl-leucyl-phenylalanine (FMLP) receptors, resulting in the inactivation of neutrophils and an inability to recognize the invading pathogens [92]. The observation that Rgps are able to strongly suppress the chemiluminescence responses of neutrophils to serum-activated zymosan [93] implies a role for gingipains in attenuating the bactericidal activities of neutrophils. Additionally, Kgp has been reported to be involved in the degradation of a C5a receptor (C5aR) on the surface of phagocytes, resulting in a suppression of neutrophil migration to infected sites [94–95].

The complement system plays a significant role in the innate immune response via three pathways: the classical, the lectin and alternative pathways. This cascade system includes about 20 glycoproteins that work in a sequence to form the membrane attack complex (MAC), which is a complex composed of C5b, C6, C7, C8 and C9 factors [96]. Upon activation, the MAC aggregates to form transmembrane channels on the bacterial cell surface, leading to the death of the bacteria caused by ruptured membranes [96]. Gingipains have been found to be able to degrade C3, C4 and C5, preventing the activation of C5b and the subsequent formation of MAC [34].

Bleeding on probing is an important clinical sign in the diagnosis of periodontitis [3, 97]. RgpA has been found to stimulate the activation of thrombin in the periodontal lesion by directly activating factor X, factor IX, prothrombin and protein C in the coagulation cascade [34]. The uncontrolled activation of thrombin stimulates the production of prostaglandin, platelet-activating factor and interleukin-1,

which are major factors in tissue inflammation and destruction [5]. Additionally, Kgp digests fibrinogen and fibrin efficiently, which are major components of plasma and responsible for blood coagulation [34, 53]. Hence, by working in unison the gingipains contribute to bleeding by deregulating the host coagulation system.

Rgps are also able to activate prekallikrein to generate kallikrein, a protease which is responsible for the formation of bradykinin from high molecular weight kininogen [5, 34]. Additionally, Rgps and Kgp together are able to release kallidin directly from kininogen [34]. Bradykinin and kallidin, also known as kinins, enhance the vascular permeability and the generation of gingival crevicular fluid [5, 34]. They also activate prostaglandin synthesis in periodontal-ligament cells and osteoblasts, which leads to alveolar bone resorption [34].

P. gingivalis is also known to evade the host immune system by deregulating the cytokine network [98–99], which comprises signalling proteins and their receptors. Gingipains have been shown to degrade and inactivate many of the components of the cytokine network including the interleukins (IL), IL-1 β , IL-4, IL-6, IL-8 and IL-12, the IL-6 receptor, interferon- α , tumour necrosis factor- α (TNF- α), CD4, CD8, CD14 and CD54 [34]. This results in the disruption of critical communications within the immune system. By deregulating the host cytokine network, the bacterium is able to evade destruction by immune cells.

Gingipains in tissue matrix destruction

It is well known that gingipains are able to degrade a number of extracellular matrix proteins *in vitro*, including laminin, fibronectin, collagen types III, IV and V, and possibly collagen type I which is the major component of collagen fibers in gingival connecting tissue [5, 100]. However, it has been pointed out that the degradation of matrix proteins by gingipain proteolytic activity alone is insufficient to cause the destruction of the gingival connective tissue. For this to occur, the endogenous matrix-degrading metalloproteinases also make a significant contribution [5, 34].

Matrix metalloproteinases (MMPs) are a group of zinc- and calcium-dependent endopeptidases produced as zymogens (pro-MMPs) by many cells including dendritic cells, periodontal ligament cells, gingival fibroblasts and the oral mucosa [34]. The pro-MMPs are activated by the removal of the pro-fragment, usually through proteolysis regulated by the host. The activities of MMPs are inhibited by tissue inhibitor metalloproteinases (TIMPs). It has been reported that Rgp is able to up regulate the expression of MMP-1 from human gingival fibroblasts [101] and to activate MMPs directly [102–103]. Such actions could result in an imbalance between MMPs and TIMPs, which would contribute to the destruction of the matrix tissue.

Neutrophils, which are attracted by host inflammatory responses, infiltrate the periodontal tissue and release elastase, cathepsin G, protease 3, MMP-8 and MMP-9 [34]. Such enzymes are normally repressed by inhibitors pro-

duced by the host. However, once again it has been reported that gingipains are able to proteolytically inactivate these inhibitors [104–107], leading to destructive proteolysis of the periodontal tissue.

The modular structure of the HA regions of gingipains

The Protein families (Pfam) database

Sequence alignments between RgpA, Kgp and HagA have shown that some peptides are highly conserved in their HA regions. The sequence similarities have been thoroughly analysed in the Pfam database and an alternative domain architecture has been predicted for gingipains. The Pfam database (<http://pfam.sanger.ac.uk/>) is a large and comprehensive collection of protein families constructed using multiple sequence alignments and profile hidden Markov models [108]. Pfam-A contains manually curated families of a high quality. The construction of each Pfam-A family requires an annotation, a seed alignment, a profile hidden Markov model and a multiple-sequence alignment. The seed alignment is made by aligning a set of sequence representatives for which there is reliable evidence that they are family members and that the sequences are not very similar. The seed alignment is then used to build up a profile hidden Markov model. This profile hidden Markov model is then used to search homologues in any sequence database. Subsequently, a full alignment is made by aligning the sequence homologues with the profile hidden Markov model. A domain family is then assigned for all of the aligned sequences. The domain families with a similar fold or predicted fold are included in a superfamily or clan. For each protein sequence entry, a schematic domain organisation is made up based on the predicted domains.

Sequence alignments between RgpA, Kgp and HagA have shown that some peptides are highly conserved in their haemagglutinin/adhesin regions. The sequence similarities have been thoroughly analysed and presented in the Pfam database and a modular architecture has been predicted for the HA regions of gingipains.

The cleaved adhesin domain family

There are presently a total of 102 entries in the cleaved adhesin family (PF07675). These cleaved adhesin domains are found in 47 protein entries from *P. gingivalis*, *Psychroflexus torques*, *Herpetosiphon aurantiacus*, *Candidatus Cloacamonas acidaminovorans* and *Flavobacterium bacterium* BAL38, with 20 unique domain organizations. Within the 30 protein sequences from *P. gingivalis*, 15 are full length gingipains (RgpA and Kgp) and three are HagA proteins from strain W83, 381 and ATC33277 containing eight or ten cleaved adhesin domains. The alternative domain structure model is presented for RgpA, Kgp and HagA in Fig. 4 and the set of Pfam models chosen are represen-

tative of all of the adhesin modules found in disease related strains of *P. gingivalis*. In order to refer to the cleaved adhesin domain model, we have named the modules K1, K2 and K3 in Kgp, R1 and R2 in RgpA and A1 to A10 in HagA.

The domain structures of full length RgpA or Kgp predicted by Pfam comprise two regions: a catalytic domain region and a cleaved adhesin domain region. The catalytic domain regions of RgpA and Kgp have a similar construction, composed of one signal peptide, one C25 propeptide, one C25 peptidase domain and one peptidase C terminal Ig-like domain. The cleaved adhesin domain regions of RgpA and a Kgp HG66 contain two cleaved adhesin domains and two DUF2436 domains of unknown function which flank the two cleaved adhesins, while three cleaved

adhesin domains and one DUF2436 domain are found in all other Kgps (Fig. 4).

In a comparison of the Pfam predicted domain model with the gingipain domain structure hypothesis (Fig. 2) it is evident that the domain organizations in the catalytic domain regions of the two models are basically the same, however the domain boundaries of the HA regions differ substantially. K2 and R2 domains, containing 172 residues in each, have the same N-terminal boundaries as those defined for Kgp15 and Rgp15. K1 and R1 domains are located in the C-terminal regions of Kgp39 or Rgp44, respectively, while the K3 domain and the C-terminal DUF2436 domain are located at the centre of the Rgp17–Rgp27 and Kgp44 regions and share an identical N-terminus with Kgp13.

Fig. 4. Domain structures of RgpA, Kgp and HagA. RgpA is from strain W50, Kgps are from strains HG66, 381, W83 and W83v. HagA is from strain 381. The domains are presented as colored boxes and residues at the N-terminus and C-terminus for each domain are labeled at the top of each box. The cleaved adhesin domains are assigned to be R1 and R2 in RgpA, K1, K2, K3, K1* and K3* in Kgps and A1–A10 in HagA. The green, yellow and blue bars represent the flanking domain locations of three adhesion binding motifs ABM1, ABM2 and ABM3. The regions labelled as DUF2436 are possibly domains of unknown function

Comparison of the cleaved adhesin modules found in gingipains

To eliminate degenerate sequences in the cleaved adhesin domain family from consideration, nine sequences from *P. gingivalis* were chosen as seeds to represent the sequence variety. The multiple amino acid sequence alignment including strain specific variations for all of the 110 sequences can be viewed at <http://pfam.sanger.ac.uk>.

Sequence alignments of the sequences of 31 cleaved adhesin domains from RgpA, Kgp and HagA from strains HG66, 381, W83, W83v and W50 have shown that they form five tight homology groups (labelled in five different colours and defined as groups 1–5 and shown in Fig. 5). K2 is conserved in all Kgps and K1 and R1 are identical in all strains.

The sequences of K2, R2, A3, A5, A7 and A10 are 97–100% identical. Kgp W83v effectively has a K3–K2–K3 combination as the K1* and K3 domains of this protein are 94% identical. K3, A8 of HagA1 and A10 of HagA2 from W83 strains are almost identical, and K1*, A2, A4, A6 and A8 of HagA2 are similar (90% identical) to K3. The K3* domain is the most different in the set and sits alone being at best ~40% identical to modules in other groups. The sequence identities between members of the five groups are in a range of 30–74%. These close relationships indicate that this set of groups of modules is a result of gene duplication. The high level of conservation of amino acid sequence in these homologues may also indicate functional conservation in addition to the predicted conservation of structure.

Fig. 5. Pairwise sequence comparisons of 31 cleaved adhesin domains from RgpA, Kgp and HagA from a complete family multiple sequence alignment. The conservation of amino acid sequences in RgpA, Kgp and the HagA proteins is shown by colour groups. Each colour group represents regions of amino acid sequence in *P. gingivalis* proteins which in pairwise alignments are more than 96% identical. The number on the line linking domains of different colour groups is the percentage of sequence identity between each pair. The five groups of adhesins are defined here as follows: Group 1=K1 and R1; Group 2=A2, A4, A6, A8, A10, K1* and K3; Group 3=A1; Group 4=A3, A5, A7, A9, K2 and R2; and Group 5=K3*. The multiple sequence alignments are provided as an Appendix in supplementary material

The galactose-binding domain-like superfamily

The hidden Markov model predicts that the cleaved adhesin domain family belongs to the galactose-binding domain-like superfamily (GBD) (PFam Clan CL0202). Observed structures of the family members in the GBD superfamily indicate that they share a β -sandwich jelly roll topology despite sequence similarities between many of them being insignificant. Twenty-five protein families are found in this clan representing a diversity of biological functions. For example, the Ephrin receptor ligand binding domain family are involved in protein–protein interactions and bind to ephrins. The MAM domain family proteins are involved in cell–cell interactions, while carbohydrate binding module families CBM4–9, CBM6, CBM11 and CBM17–28 (as number by the CAZY database) are sub-domains of glycoside hydrolases, many of which are involved in plant cell wall degradation. It is not known whether the cleaved adhesin domain family is functionally related to any of the other families or if it is only structurally similar.

Although the haemagglutinin/adhesin regions of gingipains have been related to the virulence of *P. gingivalis* in colonization, haemagglutination, haemolysis, and haem acquisition, the exact mechanisms by which gingipains interact with the host remain unclear [34]. Biochemical studies based upon exploiting the predicted modular HA model combined with further structural studies of these modules may be the approach that will lead us to a comprehensive understanding of the mechanisms of gingipain function.

The structures of cleaved adhesin modules

The Pfam predicted domain structures of the HA regions of gingipains have provided new information on the probable domain boundaries and of possible similarities between the cleaved adhesin domains. Based on these models, a number of recombinant protein constructs suitable for structural studies have been the subject of close investigation. Several of these recombinant protein constructs can be produced as soluble proteins *in vitro* (Nan Li personal communication). The crystal structure of the W83 Kgp K2 (Group 4) module has recently been reported and refined at high resolution [109] (Fig. 6A). As predicted the K2 module folds as a β -sandwich jelly roll globular domain. This module shares a significant conservation of sequence with modular sections found in each of the Kgp, RgpA and HagA proteins expressed by this and other strains (Fig. 5). More recently the crystal structure of the K3 (Group 2) module of W83 Kgp refined at a resolution of 1.6 Å has been deposited at the PDB (code: 3M1H) (Fig. 6B). The K3 fold is also a β -sandwich module and as reported for K2 it is also stabilised by two Ca^{2+} ions (Fig. 6). It is not uncommon that protein domains such as K2 (Group 4) and K3 (Group 2) with only 33% amino acid sequence identity are observed to be such close structural homologues. This fold is intimately related to other families of the galactose-bind-

ing domain-like superfamily. Given the sequence identity of 71% between K3 and K1 (Group 1) in strain W83, the fold of the K1 module in Kgp will be almost identical. It can therefore be concluded that HA region of Kgp W83 is composed of a region unknown structure (the so-called DUF2436 domain) followed by a three repeats in tandem of homologous cleaved adhesin modules. These four domains are flanked and linked by sequences that represent adhesion binding motifs (ABM1, ABM2 and ABM3) (Fig. 4).

As three of the cleaved adhesin domain group structures are not known (Groups 1, 2, 4) it is highly likely, given the known sequence similarities and identities (Fig. 5), that all the members of the *P. gingivalis* cleaved adhesin Groups 1–5 (Fig. 5) are indeed homologues.

Unravelling the functions of cleaved adhesin modules

As it has only recently become apparent that the HA regions in gingipains are composed of tandem repeats of discrete adhesin modules, functional studies of these discrete gingipain components have only just commenced. Possible functions of the K2 module have been investigated. Purified Kgp and the recombinant K2 module induce the *in vitro* haemolysis of erythrocytes with similar efficiencies. The observed dose-dependent responses can be augmented by the blocking of anion transport [109]. Further, cysteine-activated arginine gingipain RgpB, which degrades glycophorin A, sensitizes erythrocytes to the haemolytic effect of K2. Cleaved K2, similar to that found in extracted Kgp preparations, lacks the haemolytic activity indicating that autolysis of Kgp may be a staged process which is artificially enhanced by extraction of the protein [109]. This hypothesis assumes a function for K2 in the role of Kgp that enables the porphyrin auxotroph to acquire essential haem from erythrocytes.

Any likely hypothesis that addresses the basis of the polyfunctional nature of the HA regions and gingipains is to be found in the different and cooperative binding properties of individual adhesin modules. Specific sets of modules may confer particular binding activities on gingipains, thereby targeting the catalytic domains to specific host proteins. Further, it might be that some sets of these modules act synergistically to bind to those protein targets. How these multi-domain proteins may be organized in the context of the bacterial surface, perhaps even associated with other partner proteins in gingipain complexes, is yet to be determined.

Future investigations of the specific recognition of host protein/carbohydrate targets by the cleaved modular adhesins found in the HA regions and related HagA proteins will contribute to a greater understanding of complex gingipain functions and their role in *P. gingivalis* biology. For the development of clinical applications, more knowledge is required about the specific role of gingipains as virulence factors, and particularly of their interactions with the host immune system.

Fig. 6. Ribbon representations of the crystal structures of cleaved adhesin modules coloured from N to C termini with a colour ramp from blue to red. **A.** K2 from strain W83 (PDB code: 3KM5) and **B.** K3 from strain W83 (PDB code: 3M1H). **C.** A projection of the three dimensional fold topology of the K3 structure. The β -strands in the two β -sheets which make up the jelly roll fold are coloured green and yellow, respectively, with two small β -strands located in the loop region coloured blue. As indicated by a visual inspection of A and B the topology of K3 as shown is almost identical to that reported for K2 [109]

Acknowledgement

This work was supported by a grant (LP0562660) from the Australian Research Council.

References

- Williams RC: Periodontal disease. *N Engl J Med* 322, 373–382 (1990)
- Kinane DF: Causation and pathogenesis of periodontal disease. *Periodontol* 2000 25, 8–20 (2001)
- Flemmig TF: Periodontitis. *Ann Periodontol* 4, 32–38 (1999)
- Lindhe J, Ranney R, Lamster I: Consensus report: Chronic periodontitis. *Annals of Periodontology* 4, 1 (1999)
- Potempa J, Banbula A, Travis J: Role of bacterial proteinases in matrix destruction and modulation of host responses. *Periodontol* 2000 24, 153–192 (2000)
- Humphrey LL, Fu R, Buckley DI, Freeman M, Helfand M: Periodontal disease and coronary heart disease incidence: a systematic review and meta-analysis. *J Gen Intern Med* 23, 2079–2086 (2008)
- Oliver RC, Tervonen T: Periodontitis and tooth loss: comparing diabetics with the general population. *J Am Dent Assoc* 124, 71–76 (1993)
- Katz PP, Wirthlin MR, Jr., Szpunar SM, Selby JV, Sepe SJ, Showstack JA: Epidemiology and prevention of periodontal disease in individuals with diabetes. *Diabetes Care* 14, 375–385 (1991)
- Smolik I, Robinson D, El-Gabalawy HS: Periodontitis and rheumatoid arthritis: epidemiologic, clinical, and immunologic associations. *Compend Contin Educ Dent* 30, 188–190, 192, 194 passim; quiz 198, 210 (2009)
- de Pablo P, Chapple IL, Buckley CD, Dietrich T: Periodontitis in systemic rheumatic diseases. *Nat Rev Rheumatol* 5, 218–224 (2009)
- Kamer AR, Craig RG, Dasanayake AP, Brys M, Glodzik-Sobanska L, de Leon MJ: Inflammation and Alzheimer's disease: possible role of periodontal diseases. *Alzheimers Dement* 4, 242–250 (2008)
- Kamer AR, Craig RG, Pirraglia E, Dasanayake AP, Norman RG, Boylan RJ et al.: TNF-alpha and antibodies to periodontal bacteria discriminate between Alzheimer's disease patients and normal subjects. *J Neuroimmunol* 216, 92–97 (2009)
- Kamer AR, Dasanayake AP, Craig RG, Glodzik-Sobanska L, Bry M, de Leon MJ: Alzheimer's disease and peripheral infections: the possible contribution from periodontal infections, model and hypothesis. *J Alzheimers Dis* 13, 437–449 (2008)
- Loesche WJ: Chemotherapy of dental plaque infections. *Oral Sci Rev* 9, 65–107 (1976)
- Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL, Jr.: Microbial complexes in subgingival plaque. *J Clin Periodontol* 25, 134–144 (1998)
- Socransky SS, Haffajee AD: Periodontal microbial ecology. *Periodontol* 2000 38, 135–187 (2005)
- Moore WE, Holdeman LV, Cato EP, Smibert RM, Burmeister JA, Ranney RR: Bacteriology of moderate (chronic) periodontitis in mature adult humans. *Infect Immun* 42, 510–515 (1983)
- Holt SC, Kesavalu L, Walker S, Genco CA: Virulence factors of *Porphyromonas gingivalis*. *Periodontol* 2000 20, 168–238 (1999)
- Socransky SS, Haffajee AD: Evidence of bacterial etiology: a historical perspective. *Periodontol* 2000 5, 7–25 (1994)
- Holt SC, Ebersole J, Felton J, Brunsdold M, Kornman KS: Implantation of *Bacteroides gingivalis* in nonhuman primates initiates progression of periodontitis. *Science* 239, 55–57 (1988)
- Slots J, Listgarten MA: *Bacteroides gingivalis*, *Bacteroides intermedius* and *Actinobacillus actinomycetemcomitans* in human periodontal diseases. *J Clin Periodontol* 15, 85–93 (1988)
- van Steenberg TJ, van Winkelhoff AJ, van der Velden U, de Graaff J: Taxonomy, virulence and epidemiology of black-pigmented *Bacteroides* species in relation to oral infections. *Infection* 17, 194–196 (1989)
- Pathirana RD, O'Brien-Simpson NM, Veith PD, Riley PF, Reynolds EC: Characterization of proteinase-adhesin complexes of *Porphyromonas gingivalis*. *Microbiology* 152, 2381–2394 (2006)
- Lamont RJ, Jenkinson HF: Life below the gum line: pathogenic mechanisms of *Porphyromonas gingivalis*. *Microbiol Mol Biol Rev* 62, 1244–1263 (1998)
- Potempa J, Sroka A, Imamura T, Travis J: Gingipains, the major cysteine proteinases and virulence factors of *Porphyromonas gingivalis*: structure, function and assembly of multidomain protein complexes. *Curr Protein Pept Sci* 4, 397–407 (2003)
- Mayrand D, Holt SC: Biology of asaccharolytic black-pigmented *Bacteroides* species. *Microbiol Rev* 52, 134–152 (1988)
- Kusaba A, Ansai T, Akifusa S, Nakahigashi K, Taketani S, Inokuchi H et al.: Cloning and expression of a *Porphyromonas gingivalis* gene for protoporphyrinogen oxidase by complementation of a hemG mutant of *Escherichia coli*. *Oral Microbiol Immunol* 17, 290–295 (2002)
- Roper JM, Raux E, Brindley AA, Schubert HL, Gharbia SE, Shah HN et al.: The enigma of cobalamin (Vitamin B12) biosynthesis in *Porphyromonas gingivalis*. Identification and characterization of a functional corrin pathway. *J Biol Chem* 275, 40316–40323 (2000)
- Smalley JW, Birss AJ, Szmigielski B, Potempa J: The HA2 haemagglutinin domain of the lysine-specific gingipain (Kgp) of *Porphyromonas gingivalis* promotes micro-oxo bishaem formation from monomeric iron(III) protoporphyrin IX. *Microbiology* 152, 1839–1845 (2006)
- Naito M, Hirakawa H, Yamashita A, Ohara N, Shoji M, Yukitake H et al.: Determination of the genome sequence of *Porphyromonas gingivalis* strain ATCC 33277 and genomic comparison with strain W83 revealed extensive genome rearrangements in *P. gingivalis*. *DNA Res* 15, 215–225 (2008)
- Eichinger A, Beisel HG, Jacob U, Huber R, Medrano FJ, Banbula A et al.: Crystal structure of gingipain R: an Arg-specific bacterial cysteine proteinase with a caspase-like fold. *Embo J* 18, 5453–5462 (1999)
- Curtis MA, Kuramitsu HK, Lantz M, Macrina FL, Nakayama K, Potempa J et al.: Molecular genetics and nomenclature of proteases of *Porphyromonas gingivalis*. *J Periodontol Res* 34, 464–472 (1999)
- Potempa J, Pike R, Travis J: The multiple forms of trypsin-like activity present in various strains of *Porphyromonas gingivalis* are due to the presence of either Arg-gingipain or Lys-gingipain. *Infect Immun* 63, 1176–1182 (1995)
- Guo Y, Nguyen KA, Potempa J: Dichotomy of gingipains action as virulence factors: from cleaving substrates with the

- precision of a surgeon's knife to a meat chopper-like brutal degradation of proteins. *Periodontol* 2000 54, 15–44 (2010)
35. Okamoto K, Kadowaki T, Nakayama K, Yamamoto K: Cloning and sequencing of the gene encoding a novel lysine-specific cysteine proteinase (Lys-gingipain) in *Porphyromonas gingivalis*: structural relationship with the arginine-specific cysteine proteinase (Arg-gingipain). *J Biochem* 120, 398–406 (1996)
 36. Pavloff N, Pemberton PA, Potempa J, Chen WC, Pike RN, Prochazka V et al.: Molecular cloning and characterization of *Porphyromonas gingivalis* lysine-specific gingipain. A new member of an emerging family of pathogenic bacterial cysteine proteinases. *J Biol Chem* 272, 1595–1600 (1997)
 37. Barkocy-Gallagher GA, Han N, Patti JM, Whitlock J, Progulski-Fox A, Lantz MS: Analysis of the *prtP* gene encoding porphypain, a cysteine proteinase of *Porphyromonas gingivalis*. *J Bacteriol* 178, 2734–2741 (1996)
 38. Slakeski N, Bhogal PS, O'Brien-Simpson NM, Reynolds EC: Characterization of a second cell-associated Arg-specific cysteine proteinase of *Porphyromonas gingivalis* and identification of an adhesin-binding motif involved in association of the *prtR* and *prtK* proteinases and adhesins into large complexes. *Microbiology* 144, 1583–1592 (1998)
 39. Okamoto K, Nakayama K, Kadowaki T, Abe N, Ratnayake DB, Yamamoto K: Involvement of a lysine-specific cysteine proteinase in hemoglobin adsorption and haem accumulation by *Porphyromonas gingivalis*. *J Biol Chem* 273, 21225–21231 (1998)
 40. Pavloff N, Potempa J, Pike RN, Prochazka V, Kiefer MC, Travis J et al.: Molecular cloning and structural characterization of the Arg-gingipain proteinase of *Porphyromonas gingivalis*. Biosynthesis as a proteinase-adhesin polypeptide. *J Biol Chem* 270, 1007–1010 (1995)
 41. Nakayama K: Domain-specific rearrangement between the two Arg-gingipain-encoding genes in *Porphyromonas gingivalis*: possible involvement of nonreciprocal recombination. *Microbiol Immunol* 41, 185–196 (1997)
 42. Bhogal PS, Slakeski N, Reynolds EC: A cell-associated protein complex of *Porphyromonas gingivalis* W50 composed of Arg- and Lys-specific cysteine proteinases and adhesins. *Microbiology* 143, 2485–2495 (1997)
 43. Slakeski N, Cleal SM, Bhogal PS, Reynolds EC: Characterization of a *Porphyromonas gingivalis* gene *prtK* that encodes a lysine-specific cysteine proteinase and three sequence-related adhesins. *Oral Microbiol Immunol* 14, 92–97 (1999)
 44. Veith PD, Talbo GH, Slakeski N, Dashper SG, Moore C, Paolini RA et al.: Major outer membrane proteins and proteolytic processing of RgpA and Kgp of *Porphyromonas gingivalis* W50. *Biochem J* 363, 105–115 (2002)
 45. DeCarlo AA, Paramasvaran M, Yun PL, Collyer C, Hunter N: Porphyrin-mediated binding to hemoglobin by the HA2 domain of cysteine proteinases (gingipains) and haemagglutinins from the periodontal pathogen *Porphyromonas gingivalis*. *J Bacteriol* 181, 3784–3791 (1999)
 46. Pike R, McGraw W, Potempa J, Travis J: Lysine- and arginine-specific proteinases from *Porphyromonas gingivalis*. Isolation, characterization, and evidence for the existence of complexes with haemagglutinins. *J Biol Chem* 269, 406–411 (1994)
 47. Allaker RP, Aduse-Opoku J, Batten JE, Curtis MA: Natural variation within the principal arginine-specific protease gene, *prpR1*, of *Porphyromonas gingivalis*. *Oral Microbiol Immunol* 12, 298–302 (1997)
 48. Mikolajczyk-Pawlinska J, Kordula T, Pavloff N, Pemberton PA, Chen WC, Travis J et al.: Genetic variation of *Porphyromonas gingivalis* genes encoding gingipains, cysteine proteinases with arginine or lysine specificity. *Biol Chem* 379, 205–211 (1998)
 49. O'Brien-Simpson NM, Veith PD, Dashper SG, Reynolds EC: *Porphyromonas gingivalis* gingipains: the molecular teeth of a microbial vampire. *Curr Protein Pept Sci* 4, 409–426 (2003)
 50. Nadkarni MA, Nguyen KA, Chapple CC, DeCarlo AA, Jacques NA, Hunter N: Distribution of *Porphyromonas gingivalis* biotypes defined by alleles of the *kgp* (Lys-gingipain) gene. *J Clin Microbiol* 42, 3873–3876 (2004)
 51. Nguyen KA, Travis J, Potempa J: Does the importance of the C-terminal residues in the maturation of RgpB from *Porphyromonas gingivalis* reveal a novel mechanism for protein export in a subgroup of Gram-negative bacteria? *J Bacteriol* 189, 833–843 (2007)
 52. Han N, Whitlock J, Progulski-Fox A: The haemagglutinin gene A (*hagA*) of *Porphyromonas gingivalis* 381 contains four large, contiguous, direct repeats. *Infect Immun* 64, 4000–4007 (1996)
 53. Fitzpatrick RE, Wijeyewickrema LC, Pike RN: The gingipains: scissors and glue of the periodontal pathogen, *Porphyromonas gingivalis*. *Future Microbiol* 4, 471–487 (2009)
 54. Nakayama K, Ratnayake DB, Tsukuba T, Kadowaki T, Yamamoto K, Fujimura S: Haemoglobin receptor protein is intragenically encoded by the cysteine proteinase-encoding genes and the haemagglutinin-encoding gene of *Porphyromonas gingivalis*. *Mol Microbiol* 27, 51–61 (1998)
 55. Paramasvaran M, Nguyen K-A, Caldon E, McDonald JA, Najdi S, Gonzaga G et al.: Porphyrin-mediated cell surface haem capture from hemoglobin by *Porphyromonas gingivalis*. *J Bacteriol* 185, 2528–2537 (2003)
 56. DeCarlo AA, Nadkarni M, Paramasvaran M, Yun PW, Collyer CA, Hunter N: Serum antibodies against the hemoglobin-binding domain (HA2) of *Porphyromonas gingivalis*. *J Periodontol Res* 39, 228–235 (2004)
 57. Nguyen KA: Aspects of the innate and acquired immune response to the gingipains of *Porphyromonas gingivalis*. PhD thesis (2004)
 58. Okuda K, Slots J, Genco RJ: *Bacteroides gingivalis*, *Bacteroides asaccharolyticus*, and *Bacteroides melanogenicus* sub-species: cell surface morphology and adherence to erythrocytes and human bucca epithelial cells. *Current Microbiology* 6, 7–12 (1981)
 59. Sakai E, Naito M, Sato K, Hotokezaka H, Kadowaki T, Kamaguchi A et al.: Construction of recombinant haemagglutinin derived from the gingipain-encoding gene of *Porphyromonas gingivalis*, identification of its target protein on erythrocytes, and inhibition of haemagglutination by an interdomain regional peptide. *J Bacteriol* 189, 3977–3986 (2007)
 60. Okuda K, Takazoe I: Haemagglutinating activity of *Bacteroides melaninogenicus*. *Arch Oral Biol* 19, 415–416 (1974)
 61. Shah HN, Gharbia SE: Lysis of erythrocytes by the secreted cysteine proteinase of *Porphyromonas gingivalis* W83. *FEMS Microb Letters* 52, 213–217 (1989)
 62. Weinberg A, Belton CM, Park Y, Lamont RJ: Role of fimbriae in *Porphyromonas gingivalis* invasion of gingival epithelial cells. *Infect Immun* 65, 313–316 (1997)
 63. Okuda K, Kato T: Haemagglutinating activity of lipopolysaccharides from subgingival plaque bacteria. *Infect Immun* 55, 3192–3196 (1987)

64. Du L, Pellen-Mussi P, Chandad F, Mouton C, Bonnaure-Mallet M: Fimbriae and the haemagglutinating adhesin HA-Ag2 mediate adhesion of *Porphyromonas gingivalis* to epithelial cells. *Infect Immun* 65, 3875–3881 (1997)
65. Ogawa T, Uchida H, Amino K: Immunobiological activities of chemically defined lipid A from lipopolysaccharides of *Porphyromonas gingivalis*. *Microbiology* 140, 1209–1216 (1994)
66. Shi Y, Ratnayake DB, Okamoto K, Abe N, Yamamoto K, Nakayama K: Genetic analyses of proteolysis, hemoglobin binding, and haemagglutination of *Porphyromonas gingivalis*. Construction of mutants with a combination of *rgpA*, *rgpB*, *kgp*, and *hagA*. *J Biol Chem* 274, 17955–17960 (1999)
67. Booth V, Ashley FP, Lehner T: Passive immunization with monoclonal antibodies against *Porphyromonas gingivalis* in patients with periodontitis. *Infect Immun* 64, 422–427 (1996)
68. Curtis MA, Aduse-Opoku J, Slaney JM, Rangarajan M, Booth V, Cridland J et al.: Characterization of an adherence and antigenic determinant of the ArgI protease of *Porphyromonas gingivalis* which is present on multiple gene products. *Infect Immun* 64, 2532–2539 (1996)
69. Booth V, Lehner T: Characterization of the *Porphyromonas gingivalis* antigen recognized by a monoclonal antibody which prevents colonization by the organism. *J Periodontal Res* 32, 54–60 (1997)
70. Kelly CG, Booth V, Kendal H, Slaney JM, Curtis MA, Lehner T: The relationship between colonization and haemagglutination inhibiting and B cell epitopes of *Porphyromonas gingivalis*. *Clin Exp Immunol* 110, 285–291 (1997)
71. Shibata Y, Hayakawa M, Takiguchi H, Shiroza T, Abiko Y: Determination and characterization of the haemagglutinin-associated short motifs found in *Porphyromonas gingivalis* multiple gene products. *J Biol Chem* 274, 5012–5020 (1999)
72. Pike RN, Potempa J, McGraw W, Coetzer TH, Travis J: Characterization of the binding activities of proteinase-adhesin complexes from *Porphyromonas gingivalis*. *J Bacteriol* 178, 2876–2882 (1996)
73. Signas C, Raucci G, Jonsson K, Lindgren PE, Anantharamiah GM, Hook M et al.: Nucleotide sequence of the gene for a fibronectin-binding protein from *Staphylococcus aureus*: use of this peptide sequence in the synthesis of biologically active peptides. *Proc Natl Acad Sci USA* 86, 699–703 (1989)
74. Chu L, Bramanti TE, Ebersole JL, Holt SC: Haemolytic activity in the periodontopathogen *Porphyromonas gingivalis*: kinetics of enzyme release and localization. *Infect Immun* 59, 1932–1940 (1991)
75. Lewis JP, Dawson JA, Hannis JC, Muddiman D, Macrina FL: Hemoglobinas activity of the lysine gingipain protease (Kgp) of *Porphyromonas gingivalis* W83. *J Bacteriol* 181, 4905–4913 (1999)
76. Hayashi H, Nagata A, Hinode D, Sato M, Nakamura R: Survey of a receptor protein in human erythrocytes for haemagglutinin of *Porphyromonas gingivalis*. *Oral Microbiol Immunol* 7, 204–211 (1992)
77. Bramanti TE, Holt SC: Roles of porphyrins and host iron transport proteins in regulation of growth of *Porphyromonas gingivalis* W50. *J Bacteriol* 173, 7330–7339 (1991)
78. Sroka A, Sztukowska M, Potempa J, Travis J, Genco CA: Degradation of host haem proteins by lysine- and arginine-specific cysteine proteinases (gingipains) of *Porphyromonas gingivalis*. *J Bacteriol* 183, 5609–5616 (2001)
79. Shizukuishi S, Tazaki K, Inoshita E, Kataoka K, Hanioka T, Amano A: Effect of concentration of compounds containing iron on the growth of *Porphyromonas gingivalis*. *FEMS Microbiol Lett* 131, 313–317 (1995)
80. Shah HN, Bonnett R, Mateen B, Williams RA: The porphyrin pigmentation of subspecies of *Bacteroides melaninogenicus*. *Biochem J* 180, 45–50 (1979)
81. Gibbons RJ, Macdonald JB: Haemin and Vitamin K compounds as required factors for the cultivation of certain strains of *Bacteroides melaninogenicus*. *J Bacteriol* 80, 164–170 (1960)
82. Olczak T, Dixon DW, Genco CA: Binding specificity of the *Porphyromonas gingivalis* haem and hemoglobin receptor HmuR, gingipain K, and gingipain R1 for haem, porphyrins, and metalloporphyrins. *J Bacteriol* 183, 5599–5608 (2001)
83. Fujimura S, Shibata Y, Hirai K, Nakamura T: Binding of hemoglobin to the envelope of *Porphyromonas gingivalis* and isolation of the hemoglobin-binding protein. *Infect Immun* 64, 2339–2342 (1996)
84. Smalley JW, Birss AJ, Silver J: The periodontal pathogen *Porphyromonas gingivalis* harnesses the chemistry of the mu-oxo bishaem of iron protoporphyrin IX to protect against hydrogen peroxide. *FEMS Microbiol Lett* 183, 159–164 (2000)
85. Smalley JW, Thomas MF, Birss AJ, Withnall R, Silver J: A combination of both arginine- and lysine-specific gingipain activity of *Porphyromonas gingivalis* is necessary for the generation of the micro-oxo bishaem-containing pigment from haemoglobin. *Biochem J* 379, 833–840 (2004)
86. Smalley JW, Birss AJ, Szmigielski B, Potempa J: Sequential action of R- and K-specific gingipains of *Porphyromonas gingivalis* in the generation of the haem-containing pigment from oxyhaemoglobin. *Arch Biochem Biophys* 465, 44–49 (2007)
87. Gao JL, Nguyen KA, Hunter N: Characterization of a hemophore-like protein from *Porphyromonas gingivalis*. *J Biol Chem* 285, 40028–40038 (2010)
88. Oda H, Saiki K, Numabe Y, Konishi K: Effect of gamma-immunoglobulin on the asaccharolytic growth of *Porphyromonas gingivalis*. *J Periodontal Res* 42, 438–442 (2007)
89. Grenier D, Roy S, Chandad F, Plamondon P, Yoshioka M, Nakayama K et al.: Effect of inactivation of the Arg- and/or Lys-gingipain gene on selected virulence and physiological properties of *Porphyromonas gingivalis*. *Infect Immun* 71, 4742–4748 (2003)
90. Kadowaki T, Nakayama K, Yoshimura F, Okamoto K, Abe N, Yamamoto K: Arg-gingipain acts as a major processing enzyme for various cell surface proteins in *Porphyromonas gingivalis*. *J Biol Chem* 273, 29072–29076 (1998)
91. Diamond G, Beckloff N, Ryan LK: Host defense peptides in the oral cavity and the lung: similarities and differences. *J Dent Res* 87, 915–927 (2008)
92. Lala A, Amano A, Sojar HT, Radel SJ, De Nardin E: *Porphyromonas gingivalis* trypsin-like protease: a possible natural ligand for the neutrophil formyl peptide receptor. *Biochem Biophys Res Commun* 199, 1489–1496 (1994)
93. Kadowaki T, Yoneda M, Okamoto K, Maeda K, Yamamoto K: Purification and characterization of a novel arginine-specific cysteine proteinase (argingipain) involved in the pathogenesis of periodontal disease from the culture supernatant of *Porphyromonas gingivalis*. *J Biol Chem* 269, 21371–21378 (1994)
94. Jagels MA, Ember JA, Travis J, Potempa J, Pike R, Hugli TE: Cleavage of the human C5A receptor by proteinases derived from *Porphyromonas gingivalis*: cleavage of leukocyte C5a receptor. *Adv Exp Med Biol* 389, 155–164 (1996)

95. Jagels MA, Travis J, Potempa J, Pike R, Hugli TE: Proteolytic inactivation of the leukocyte C5a receptor by proteinases derived from *Porphyromonas gingivalis*. *Infection & Immunity* 64, 1984–1991 (1996)
96. Peitsch MC, Tschopp J: Assembly of macromolecular pores by immune defense systems. *Curr Opin Cell Biol* 3, 710–716 (1991)
97. Page RC, Schroeder HE: Pathogenesis of inflammatory periodontal disease. A summary of current work. *Lab Invest* 34, 235–249 (1976)
98. Kjeldsen M, Holmstrup P, Bendtzen K: Marginal periodontitis and cytokines: a review of the literature. *J Periodontol* 64, 1013–1022 (1993)
99. Wilson M, Seymour R, Henderson B: Bacterial perturbation of cytokine networks. *Infect Immun* 66, 2401–2409 (1998)
100. Takii R, Kadowaki T, Baba A, Tsukuba T, Yamamoto K: A functional virulence complex composed of gingipains, adhesins, and lipopolysaccharide shows high affinity to host cells and matrix proteins and escapes recognition by host immune systems. *Infect Immun* 73, 883–893 (2005)
101. Matsushita K, Imamura T, Tanchaoen S, Tatsuyama S, Tomikawa M, Travis J et al.: Selective inhibition of *Porphyromonas gingivalis* growth by a factor Xa inhibitor, DX-9065a. *J Periodontal Res* 41, 171–176 (2006)
102. DeCarlo AA, Jr., Windsor LJ, Bodden MK, Harber GJ, Birkedal-Hansen B, Birkedal-Hansen H: Activation and novel processing of matrix metalloproteinases by a thiol-proteinase from the oral anaerobe *Porphyromonas gingivalis*. *J Dent Res* 76, 1260–1270 (1997)
103. Grayson R, Douglas C, Heath J, Rawlinson A, Evans G: Activation of human matrix metalloproteinase 2 by gingival crevicular fluid and *Porphyromonas gingivalis*. *J Clin Periodontol* 30, 542–550 (2003)
104. Nelson D, Potempa J, Kordula T, Travis J: Purification and characterization of a novel cysteine proteinase (periodontain) from *Porphyromonas gingivalis*. Evidence for a role in the inactivation of human alpha1-proteinase inhibitor. *J Biol Chem* 274, 12245–12251 (1999)
105. Potempa J, Pike R, Travis J: Titration and mapping of the active site of cysteine proteinases from *Porphyromonas gingivalis* (gingipains) using peptidyl chloromethanes. *Biol Chem* 378, 223–230 (1997)
106. Into T, Inomata M, Kanno Y, Matsuyama T, Machigashira M, Izumi Y et al.: Arginine-specific gingipains from *Porphyromonas gingivalis* deprive protective functions of secretory leukocyte protease inhibitor in periodontal tissue. *Clin Exp Immunol* 145, 545–554 (2006)
107. Kantyka T, Latendorf T, Wiedow O, Bartels J, Glaser R, Dubin G et al.: Elafin is specifically inactivated by RgpB from *Porphyromonas gingivalis* by distinct proteolytic cleavage. *Biol Chem* 390, 1313–1320 (2009)
108. Finn RD, Tate J, Mistry J, Coghill PC, Sammut SJ, Hotz HR et al.: The Pfam protein families database. *Nucleic Acids Res* 36, D281–288 (2008)
109. Li N, Yun P, Nadkarni MA, Ghadikolaee NB, Nguyen KA, Lee M et al.: Structure determination and analysis of a haemolytic gingipain adhesin domain from *Porphyromonas gingivalis*. *Molecular Microbiology* 76, 861–873 (2010)
110. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG: Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948 (2007)

Appendix

Multiple sequence alignments of cleaved adhesin domains of Kgp, RgpA and HagA. 31 sequences of cleaved adhesin domains of Kgp W83, Kgp HG66, Kgp 381, Kgp W83v, RgpA HG66, HagA W83 and HagA 381 are aligned using the program ClustalW (Larkin et al. [110]).

The first column shows the assigned domain names corresponding to *Figures 4* and *5*, sequences are in five tight homology groups which were enclosed within boxes of different colours.

K1/R1 group (red box), K1'/K3/A2/A4/A6/A8/A10 group (blue box), A1 group (green box), K2/R2/A3/A5/A7/A9 group (orange box), K3* group (pink box).

K1	KgpW83/982-1150	-GTTLSESFENG-IPASWKTIDADGDGHGWKPG----NAPG---IAGYNS	1022
K1	KgpHG66/981-1149	-TTTLESESFENG-IPASWKTIDADGDGHGWKPG----NAPG---IAGYNS	1021
R1	RgpAW50/962-1130	-TTTLESESFENG-IPASWKTIDADGDGHGWKPG----NAPG---IAGYNS	1002
K1	Kgp381/981-1149	-TTTLESESFENG-IPASWKTIDADGDGHGWKPG----NAPG---IAGYNS	1021
A4	HagAW83/958-1130	-TTTLESESFENG-IPASWKTIDADGDGNNWTTT----PPPGGTSFAGHNS	1001
A6	HagAW83/1410-1582	-TTTLESESFENG-IPASWKTIDADGDGNNWTTT----PPPGGTSFAGHNS	1453
A6	HagA381/1418-1590	-TTTLESESFENG-IPASWKTIDADGDGNNWTTT----PPPGGTSFAGHNS	1461
A4	HagA381/962-1134	-TTTLESESFENG-IPASWKTIDADGDGNNWTTT----PPPGGTSFAGHNS	1005
A2	HagA381/506-678	-TTTLESESFENG-IPASWKTIDADGDGNNWTTT----PPPGGTSFAGHNS	549
A8	HagA381/1874-2046	-TTTLESESFENG-IPASWKTIDADGDGNNWTTT----PPPGGTSFAGHNS	1917
A2	HagAW83/506-678	-TTTLESESFENG-IPASWKTIDADGDGNNWTTT----PPPGGSSFAGHNS	549
K1*	KgpW83v/630-802	-TTTLESESFENG-IPASWKTIDADGDGNNWTTT----PPPGGSSFAGHNS	673
K3	KgpW83v/1079-1251	-AEVLNEDFENG-IPASWKTIDADGDGNNWTTT----PPPGGSSFAGHNS	1122
K3	KgpW83/1427-1599	-AEVLNEDFENG-IPASWKTIDAGDGNWTTT----PPPGGSSFAGHNS	1470
A10	HagA381/2323-2495	-AEVLNEDFENG-IPSSWKTIDADGDGNNWTTT----PPPGGSSFAGHNS	2366
A8	HagAW83/1859-2031	-AEVLNEDFENG-IPASWKTIDADGDGNNWTTT----PPPGGSSFAGHNS	1902
A1	HagAW83/230-407	-TYVLNESFDQTPLPNGWTMIDADGDGHNWLST----INVYN--TATHTG	272
A1	HagA381/230-407	-TYVLNESFDQTPLPNGWTMIDADGDGHNWLST----INVYN--TATHTG	272
K2	KgpW83v/811-982	FTETFESSTHGEA-PAEWTTIDADGDQGWL----LSSGQLDWLTAHGG	855
R2	RgpAW50/1139-1310	FTETFESSTHGEA-PAEWTTIDADGDQGWL----LSSGQLDWLTAHGG	1183
A5	HagAW83/1139-1310	FTETFESSTHGEA-PAEWTTIDADGDQGWL----LSSGQLDWLTAHGG	1183
K2	Kgp381/1158-1329	FTETFESSTHGEA-PAEWTTIDADGDQDWL----LSSGQLDWLTAHGG	1202
A7	HagAW83/1591-1762	FTETFESSTHGEA-PAEWTTIDADGDQGWL----LSSGQLGWLTAHGG	1635
A3	HagAW83/687-858	FTETFESSTHGEA-PAEWTTIDADGDQGWL----LSSGQLGWLTAHGG	731
A5	HagA381/1143-1314	FTETFESSTHGEA-PAEWTTIDADGDQGWL----LSSGQLGWLTAHGG	1187
A7	HagA381/1599-1770	FTETFESSTHGEA-PAEWTTIDADGDQGWL----LSSGQLGWLTAHGG	1643
K2	KgpHG66/1158-1329	FTETFESSTHGEA-PAEWTTIDADGDQGWL----LSSGQLDWLTAHGG	1202
A9	HagA381/2055-2226	FTETFESSTHGEA-PAEWTTIDADGDQGWL----LSSGQLDWLTAHGG	2099
A3	HagA381/687-858	FTETFESSTHGEA-PAEWTTIDADGDQGWL----LSSGQLDWLTAHGG	731
K2	KgpW83/1159-1330	FTETFESSTHGEA-PAEWTTIDADGDQGWL----LSSGQLDWLTAHGG	1203
K3*	Kgp381/1424-1589	-GELLNEDFEGDAIPTGTALDADGDGNNWDITLNEFTRGERHVLSPRA	1472
K1	KgpW83/982-1150	NGCVYSESFGLGGIGVL--TPDNYLITPALDLPNG-GKLTFFWCAQDA-	1067
K1	KgpHG66/981-1149	NGCVYSESFGLGGIGVL--TPDNYLITPALDLPNG-GKLTFFWCAQDA-	1066
R1	RgpAW50/962-1130	NGCVYSESFGLGGIGVL--TPDNYLITPALDLPNG-GKLTFFWCAQDA-	1047
K1	Kgp381/981-1149	NGCVYSESFGLGGIGVL--TPDNYLITPALDLPNG-GKLTFFWCAQDA-	1066
A4	HagAW83/958-1130	AICVSSASY-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	1069
A6	HagAW83/1410-1582	AICVSSASY-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	1501
A6	HagA381/1418-1590	AICASSASY-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	1505
A4	HagA381/962-1134	AICASSASY-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	1049
A2	HagA381/506-678	AICASSASY-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	593
A8	HagA381/1874-2046	AICVSSASY-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	1961
A2	HagAW83/506-678	AICVSSASY-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	593
K1*	KgpW83v/630-802	AICVSSASY-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	717
K3	KgpW83v/1079-1251	AICVSSASY-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	1166
K3	KgpW83/1427-1599	AICVSSASH-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	1514
A10	HagA381/2323-2495	AICVSSASY-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	2410
A8	HagAW83/1859-2031	AICVSSASY-INFEGPQ--NPDNYLVTPELSLPNG-GTLTFWVCAQDA-	1946
A1	HagAW83/230-407	DGAMFSKSWTASSGAKIDLSPDNYLVTPKFTVPENG-KLSYWVSSQE--	318
A1	HagA381/230-407	DGAMFSKSWTASSGAKIDLSPDNYLVTPKFTVPENG-KLSYWVSSQV--	318
K2	KgpW83v /811-982	TNIVVASF--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	894
R2	RgpAW50/1139-1310	TNIVVASF--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	1222
A5	HagAW83/1139-1310	TNIVVASF--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	1222
K2	Kgp381/1158-1329	TNIVVASF--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	1241
A7	HagAW83/1591-1762	TNIVVASF--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	1674
A3	HagAW83/687-858	TNIVVASF--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	770
A5	HagA381/1143-1314	TNIVVASF--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	1226
A7	HagA381/1599-1770	TNIVVASF--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	1682
K2	KgpHG66/1158-1329	TNIVVASF--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	1241
A9	HagA381/2055-2226	TNIVVASF--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	2138
A3	HagA381/687-858	TNIVVASF--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	770
K2	KgpW83/1159-1330	SNVVSSFS--WNGMA---LNPDNYLISKDVTGATK---VKYYYAVND--	1242
K3*	Kgp381/1424-1589	SNVAISYSSLLQGEYLPNPNFLITPKVEGAKK---ITYKVGSPGLP	1518

K1	KgpW83/982-1150	NYASEHYAVYASSTGNDASNFTNALLEETI-TAKGVR---SPKAIRG-R	1111
K1	KgpHG66/981-1149	NYASEHYAVYASSTGNDASNFTNALLEETI-TAKGVR---SPEAIRG-R	1110
R1	RgpAW50/962-1130	NYASEHYAVYASSTGNDASNFTNALLEETI-TAKGVR---SPEAIRG-R	1091
K1	Kgp381/981-1149	NYASEHYAVYASSTGNDASNFTNALLEETI-TAKGVR---SPEAIRG-R	1110
A4	HagAW83/958-1130	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	1091
A6	HagAW83/1410-1582	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	1543
A6	HagA381/1418-1590	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	1551
A4	HagA381/962-1134	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	1095
A2	HagA381/506-678	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	639
A8	HagA381/1874-2046	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	2007
A2	HagAW83/506-678	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	639
K1*	KgpW83v/630-802	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	763
K3	KgpW83v/1079-1251	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	1212
K3	KgpW83/1427-1599	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	1560
A10	HagA381/2323-2495	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	2456
A8	HagAW83/1859-2031	NYASEHYAVYASSTGNDASNFNALLEEVL-TAKTVVT---APEAIRGTR	1992
A1	HagAW83/230-407	PWTNEHYGVFLSTTGNEAANFTIKLLEETLGSGKPAPM---NLVKSEGVK	365
A1	HagA381/230-407	PWTNEHYGVFLSTTGNEAANFTIKLLEETLGSDKPAPM---NLVKSEGVK	365
K2	KgpW83v/811-982	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEANGAK	943
R2	RgpAW50/1139-1310	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEANGAK	1271
A5	HagAW83/1139-1310	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEANGAK	1271
K2	Kgp381/1158-1329	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEANGAK	1290
A7	HagAW83/1591-1762	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEANGAK	1723
A3	HagAW83/687-858	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEANGAK	819
A5	HagA381/1143-1314	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEADGAK	1275
A7	HagA381/1599-1770	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEADGAK	1731
K2	KgpHG66/1158-1329	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEADGAK	1290
A9	HagA381/2055-2226	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEADGAK	2187
A3	HagA381/687-858	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEADGAK	819
K2	KgpW83/1159-1330	GFPGDHYAVMISKTGTNAGDFT-VVFEETPNGINKGARFGLSTEANGAK	1291
K3*	Kgp381/1424-1589	QWSDHYALCISKSGTAAADFE-VIFEETMTYTQ-----GGAN	1555
K1	KgpW83/982-1150	IQGTWRQKTVDLPAGT---KYVAFRHFQSTDMFYIDLDEVEI-	1150
K1	KgpHG66/981-1149	IQGTWRQKTVDLPAGT---KYVAFRHFQSTDMFYIDLDEVEI-	1149
R1	RgpAW50/962-1130	IQGTWRQKTVDLPAGT---KYVAFRHFQSTDMFYIDLDEVEI-	1130
K1	Kgp381/981-1149	IQGTWRQKTVDLPAGT---KYVAFRHFQSTDMFYIDLDEVEI-	1149
A4	HagAW83/958-1130	VQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVEI-	1130
A6	HagAW83/1410-1582	VQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVEI-	1582
A6	HagA381/1418-1590	VQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVEI-	1590
A4	HagA381/962-1134	VQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVEI-	1134
A2	HagA381/506-678	VQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVEI-	678
A8	HagA381/1874-2046	VQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVEI-	2046
A2	HagAW83/506-678	VQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVEI-	678
K1*	KgpW83v/630-802	VQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVEI-	802
K3	KgpW83v/1079-1251	AQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVVI-	1251
K3	KgpW83/1427-1599	AQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVVI-	1599
A10	HagA381/2323-2495	VQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVVI-	2495
A8	HagAW83/1859-2031	VQGTWYQKTVDLPAGT---KYVAFRHFQCTDFFWINLDDVVI-	2031
A1	HagAW83/230-407	APAPYQERTIDLSAYAGQQVYLAFRHFQCTGIFRLYLDDVAV-	407
A1	HagA381/230-407	LPAPYQERTIDLSAYAGQQVYLAFRHFNSTGIFRLYLDDVAV-	407
K2	KgpW83v/811-982	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	982
R2	RgpAW50/1139-1310	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	1310
A5	HagAW83/1139-1310	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	1310
K2	Kgp381/1158-1329	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	1329
A7	HagAW83/1591-1762	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	1762
A3	HagAW83/687-858	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	858
A5	HagA381/1143-1314	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	1314
A7	HagA381/1599-1770	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	1770
K2	KgpHG66/1158-1329	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	1329
A9	HagA381/2055-2226	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	2226
A3	HagA381/687-858	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	858
K2	KgpW83/1159-1330	PQSVWIERTVDLPAGT---KYVAFRHYNCSDLNYILLDDIQ-F	1330
K3*	Kgp381/1424-1589	-----LTREKDLPGT---KYVAFRHYNCSDLNYILLDDVVI-	1589