

Crystal structure of gingipain R: an Arg-specific bacterial cysteine proteinase with a caspase-like fold

Andreas Eichinger¹, Hans-Georg Beisel¹,
Uwe Jacob¹, Robert Huber¹,
Francisco-Javier Medrano^{1,2},
Agnieszka Banbula^{3,4}, Jan Potempa^{3,4},
Jim Travis⁴ and Wolfram Bode^{1,5}

¹Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, D-82152 Martinsried, Germany, ²Departamento de Microbiología, Centro de Investigaciones Biológicas, CSIC, Velazquez 144, E-28006, Spain, ³Institute of Molecular Biology, Jagiellonian University, Al. A. Mickiewicza 3, 31-120 Krakow, Poland and ⁴Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA

⁵Corresponding author
e-mail: bode@biochem.mpg.de

A.Eichinger and H.-G.Beisel contributed equally to this work

Gingipains are cysteine proteinases acting as key virulence factors of the bacterium *Porphyromonas gingivalis*, the major pathogen in periodontal disease. The 1.5 and 2.0 Å crystal structures of free and D-Phe–Phe–Arg-chloromethylketone-inhibited gingipain R reveal a 435-residue, single-polypeptide chain organized into a catalytic and an immunoglobulin-like domain. The catalytic domain is subdivided into two subdomains comprising four- and six-stranded β -sheets sandwiched by α -helices. Each subdomain bears topological similarities to the p20–p10 heterodimer of caspase-1. The second subdomain harbours the Cys–His catalytic diad and a nearby Glu arranged around the S1 specificity pocket, which carries an Asp residue to enforce preference for Arg-P1 residues. This gingipain R structure is an excellent template for the rational design of drugs with a potential to cure and prevent periodontitis. Here we show the binding mode of an arginine-containing inhibitor in the active-site, thus identifying major interaction sites defining a suitable pharmacophor.

Keywords: caspases/crystal structure/cysteine proteinase/periodontitis/*Porphyromonas gingivalis*

Introduction

Periodontal diseases represent infections that are associated with inflammation of the gingiva, destruction of periodontal tissue, pocket formation and alveolar bone resorption. If untreated, they eventually lead to exfoliation of teeth. In addition, severe periodontitis may predispose to more serious systemic conditions such as cardiovascular diseases and the delivery of preterm infants (Page, 1998). The severity of periodontitis correlates with the presence of specific bacteria that trigger inflammatory host responses which, together with the bacterial virulence factors, cause

the majority of tissue destruction (for references, see Genco *et al.*, 1998; Lamont and Jenkinson, 1998).

Porphyromonas gingivalis, a Gram-negative, anaerobic, black-pigmented bacterium, has been implicated as the major aetiological agent in the initiation and progression of adult periodontitis (Holt *et al.*, 1988). It possesses a repertoire of virulence factors such as proteinases, fimbriae, lectin-type adhesins and haemagglutinating factors, which enable it to colonize periodontal pockets (reviewed by Cutler *et al.*, 1995; Lamont and Jenkinson, 1998). It is generally accepted that the proteolytic enzymes of this organism play a central role in the pathogenesis of periodontitis. Although several different types of proteinases are expressed by *P.gingivalis*, cysteine proteinases with Arg–Xaa and Lys–Xaa specificity, referred to as gingipains R and K, respectively, are responsible for at least 85% of the overall proteolytic activity and are recognized as the major virulence factors of this periodontal pathogen (Potempa and Travis, 1996; Kuramitsu, 1998).

Gingipain Rs (Rgps) occur either in soluble (RgpA, HRgpA, RgpB) or in membrane-associated forms and are products of two related genes, *rgpA* and *rgpB*. The *rgpA* gene-encoded gingipains are released into the medium as the single-chain, 50 kDa RgpA proteinase or the high-molecular-mass HRgpA, which is a 95 kDa non-covalent complex of a 50 kDa catalytic domain with a haemagglutinin (adhesin) domain derived from the initial *rgpA* gene product via proteolytic processing (Pavloff *et al.*, 1995; Rangarajan *et al.*, 1997). The related *rgpB* gene lacks the coding region for the haemagglutinin domains and codes for a 507-amino-acid-residue protein (Mikolajczyk-Pawlinska *et al.*, 1998); its mature C-terminally truncated product RgpB is a single-chain protein essentially identical to RgpA, with the C-terminal part from position 363 onwards showing a significant divergence, however (Potempa *et al.*, 1998). Rgps exhibit an exclusive specificity for Arg–Xaa peptide bonds and their hydrolytic activity requires activation by reducing agents such as cysteine; they are inhibited by iodoacetamide, leupeptin and Arg-chloromethylketone inhibitors, but surprisingly also by chloromethylketones with Lys or other residues at P1 acting as irreversible covalently attached inhibitors (Nakayama, 1997; Potempa *et al.*, 1997). Alkylation identified Cys244 as the putative reactive-site residue (Nishikata and Yoshimura, 1995; Potempa *et al.*, 1998). The Rgps are stabilized by calcium (Chen *et al.*, 1992; Nakayama *et al.*, 1995) and inhibited by EDTA (Fujimura *et al.*, 1998).

Besides providing nutrients for bacterial growth, the Rgps are implicated in both the processing of pro-Kgp and other housekeeping functions (reviewed by Potempa and Travis, 1996). They have been shown to be involved in the processing of the 75 kDa major outer membrane

Table I. Data collection statistics

Spacegroup	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_12_12_1$
Dataset	pCMBA	pCMBA	pCMBA	NATI	FFRCMK
	λ_1 (remote)	λ_2 (peak)	λ_3 (edge)		
Wavelength (Å)	1.0697	1.0060	1.0100	1.5418	1.5418
Cell constants	55.45 59.33 62.64	55.52 59.25 62.63	55.75 59.49 62.84	54.67 59.92 63.21	51.93 79.92 99.82
a, b, c, α , β , γ	90.00 94.70 90.00	90.00 94.65 90.00	90.00 94.78 90.00	90.00 95.20 90.00	90.00 90.00 90.00
Resolution range (Å)	17.71–1.50	17.96–1.49	17.76–1.60	27.84–2.16	15.64–2.03
Unique reflections	62 420	63 274	50 402	19 470	27 024
Multiplicity	3.3	6.6	3.2	2.4	3.3
Complete (%) (last shell/Å)	95.3 (89.6)	95.7 (82.5)	92.3 (92.3)	85.2 (52.7)	99.0 (96.3)
	(1.58–1.50)	(1.57–1.49)	(1.68–1.60)	(2.29–2.16)	(2.14–2.03)
R_{sym} (%) (last shell)	5.5 (31.2)	7.4 (34.3)	4.9 (31.9)	12.0 (32.6)	7.8 (23.1)

$$R_{\text{sym}} = (\Sigma (|I(h)_i - \langle I(h) \rangle|) / \Sigma \langle I(h) \rangle)$$

protein and in the maturation and assembly of the endogenous *P.gingivalis* fimbriae, both of which are considered as important virulence factors (Onoe *et al.*, 1996; Kadowaki *et al.*, 1998). Furthermore, they can trigger bradykinin release through prekallikrein activation (Imamura *et al.*, 1994), degrade complement factors such as C3 and C5 resulting in the attraction of neutrophils to the gingival lesion site (reviewed by DiScripio *et al.*, 1996), inactivate tumour necrosis factor α (Calkins *et al.*, 1998) and activate coagulation factors (Imamura *et al.*, 1997).

Thus, the gingipains represent a novel and unique family of cysteine proteinases whose members play pivotal roles in the aetiology of periodontitis. However, in spite of a wealth of sequential and functional data, their structural folds and mechanism of action were completely unknown.

Results and discussion

Overall structure

The RgpB structure was first determined by multiple anomalous dispersion (MAD) techniques and refined with 1.5 Å data for a monoclinic $P2_1$ cell, with three *p*-chloromercury benzamidide (HgBA) molecules situated in the active-site and at two other surface-located Cys residues (pCMBA; Table I). With this model, the 2.16 Å structure of the corresponding ligand-free RgpB was determined (NATI); furthermore, it served to solve the 2.0 Å orthorhombic $P2_12_12_1$ structure of a covalent RgpB complex with D-Phe-Phe-Arg-chloromethylketone (FFRCMK; with one molecule per asymmetrical unit; see Table I).

The RgpB molecule has the shape of a crooked one-root 'tooth', made up of an almost spherical 'crown' of average diameter 45 Å and a 40-Å-long 'root' (Figures 1 and 2A). The crown formed by the N-terminal 351 residues represents the catalytic domain, while the root made by the last 84 residues resembles an immunoglobulin superfamily (IgSF) domain. RgpB is characterized by a high ratio of regular secondary structure elements, with the catalytic domain exhibiting the structural motif of a typical α/β protein and the IgSF domain having an all- β conformation. The seven Cys residues of RgpB are unpaired, with all but three (Cys185, the catalytic Cys244 and Cys299) exhibiting buried side chains. RgpB exhibits a low isoelectric point and a comparatively large ratio of

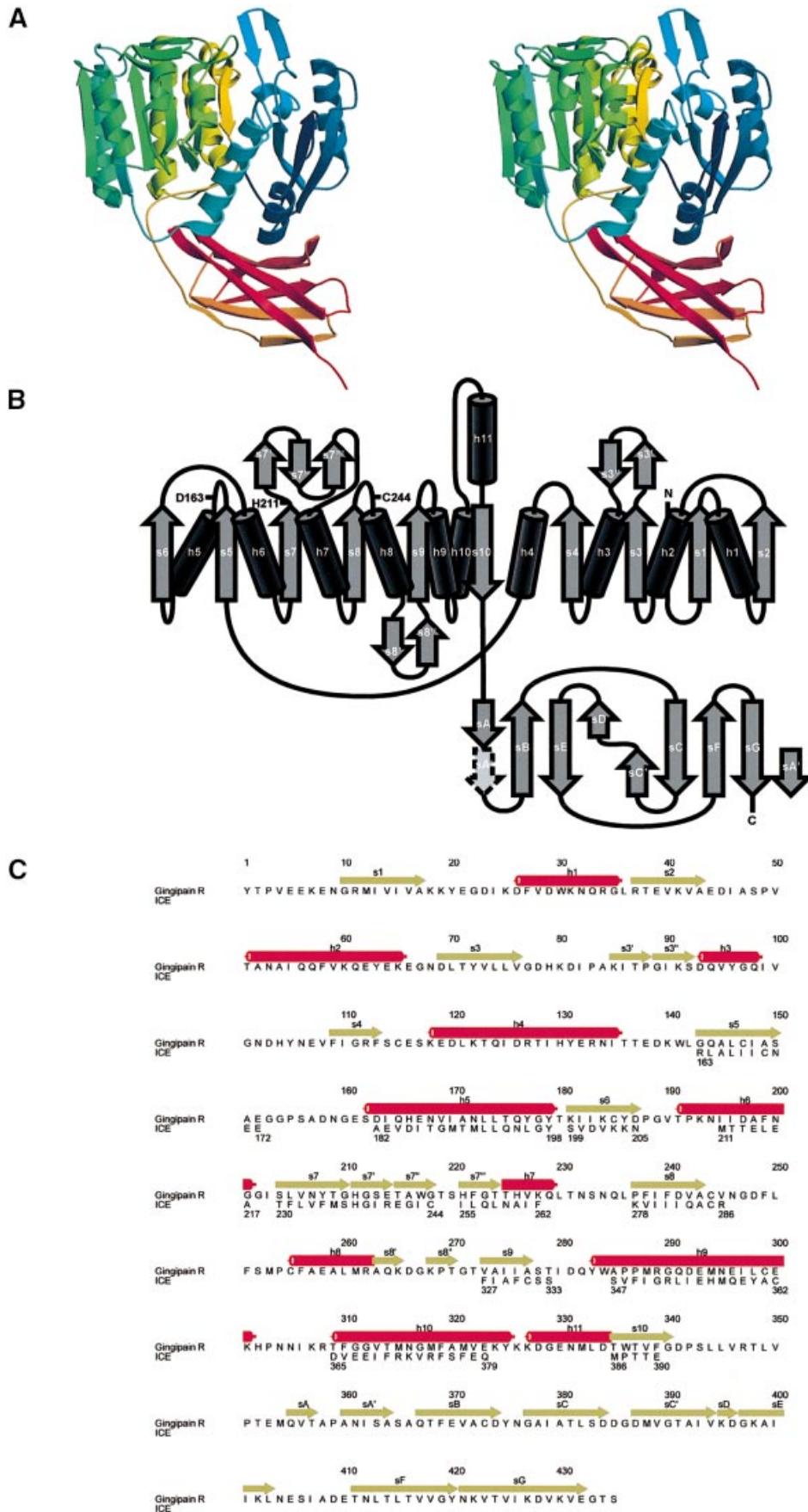


Fig. 1. Ribbon plot of RgpB (front view). The molecule consists of the catalytic domain (top) subdivided into A- (right) and B- subdomains (left), and the IgSF domain (bottom). Strands are shown as yellow arrows, helices as red spirals and the connecting segments as blue ropes. The catalytic residues on top are shown as orange stick models and the three putative calcium ions as golden spheres. The figure was prepared with MOLSCRIPT (Kraulis, 1991) and rendered with POVray (1997).

charged residues, several of which are uncompensated by opposite charges and/or buried.

The catalytic domain

The catalytic domain of RgpB can be subdivided into A- and B- subdomains. Each subdomain comprises a central β -sheet and a few additional hairpins flanked by helices on either side, as characteristic for α/β open-sheet structures. The A-subdomain sheet consisting of four parallel strands is twisted regularly, with the two outermost strands arranged at an angle of $\sim 45^\circ$ to one another (Figure 1). The flanking helices placed on both sides of this β -sheet run approximately parallel to one another, but antiparallel to the sheet strands (Figure 2A and B). The core is mainly



hydrophobic in nature, but is traversed by the Arg112–Asp93 and the Arg34–Glu132–Lys31 salt bridge/cluster.

The RgpB polypeptide chain runs over to the adjacent B-subdomain via a short surface-located segment connecting helix h4 and strand s5. Central to this subdomain is the relatively flat β -sheet made up of six strands (Figure 2). While the first five strands are arranged parallel and show almost no twist to one another, the last innermost strand (s10) runs antiparallel to the previous five strands. To the other side, strand s10 crosses the s4 edge strand of the A-subdomain sheet under the formation of two hydrogen bonds between Phe339 and Arg112. Subsequent to strand s10, the polypeptide chain enters the IgSF domain (Figure 1). The central sheet of the B-subdomain is flanked by the seven helices h5–h11, which except for h10 all run parallel to one another and antiparallel to sheet strands s6–s9, respectively. Helix h9 is special in that its N-terminal part (Tyr283–Trp284–Ala285–Pro286–Pro287), positioned close to the active centre (see Figure 3), adopts an open conformation due to the inability of Pro287 to form a hydrogen bond to the Tyr283 carbonyl group. As usual in such open β -sheet enzymes, the active centre is in a crevice outside the carboxyl end of the β -sheet, with the active-site residues His211 and Cys244 presented by loops s7–s7' and s8–h8 (and probably Glu152 on loop s5–h5; Figure 2B).

The interface between both RgpB A- and B-subdomains

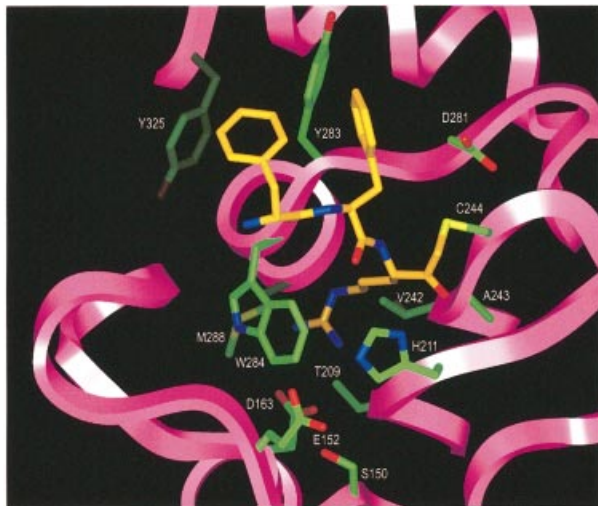


Fig. 3. Interaction of the D-Phe–L-Phe–L-Arg methylene inhibitor with the RgpB active-site. The active-site region of RgpB, besides a few important residues (green) mainly represented by the ribbon-like backbone (pink), is shown in standard orientation (obtained from the front view, Figure 1, upon a 90° rotation about a horizontal axis). The inhibitor chain (yellow stick model) covalently linked via its methylene group to Cys244 $S\gamma$ (centre, right) runs from left to right, with its Arg-P1 side chain reaching back into the S1 pocket. The imidazole side chain of His211 and the carboxylate of Glu152 are arranged on the molecular surface (bottom) opposite to Cys244. The figure was prepared with Insight II.

Fig. 2. Topology and sequence of RgpB. (A) Stereo ribbon plot of RgpB in the front view orientation of Figure 1 but colour coded by the spectral colours according to sequential order. (B) Topological diagram of RgpB. Arrows denote strands s1–s10 (catalytic domain) and sA–sH (IgSF domain), and cylinders indicate helices h1–h11 (catalytic domain). The dashed arrow (sA') emphasizes the intermediate crossing-over of IgSF strand sA toward the opposite sheet. (C) RgpB sequence from Tyr1 to Ser435. Arrows and braces indicate β -strands and helices in RgpB. A total of 118 residues of caspase-1 (Cerretti *et al.*, 1992; Thornberry *et al.*, 1992) have been aligned to the B-subdomain according to topological equivalency with caspase-1 as deposited by Rano *et al.* (1997) in the Protein Data Bank under accession code 1IBC. The numbering is for RgpB (top, SwissProt entry code P95493) and caspase-1 (bottom). The figure was prepared with ALSRIPT (Barton, 1993).

is relatively hydrophilic, containing, besides a number of buried water molecules, a cluster of charged residues (including the partially buried Arg262, the internal Glu258 and Asp78, and the surface-located Glu116) that clamp both subdomains together via a central calcium ion. At two further sites (NATI and pCMBA), or up to six other positions (FFRCMK, with additional zinc ions in the crystallization buffer) on the surface of the catalytic

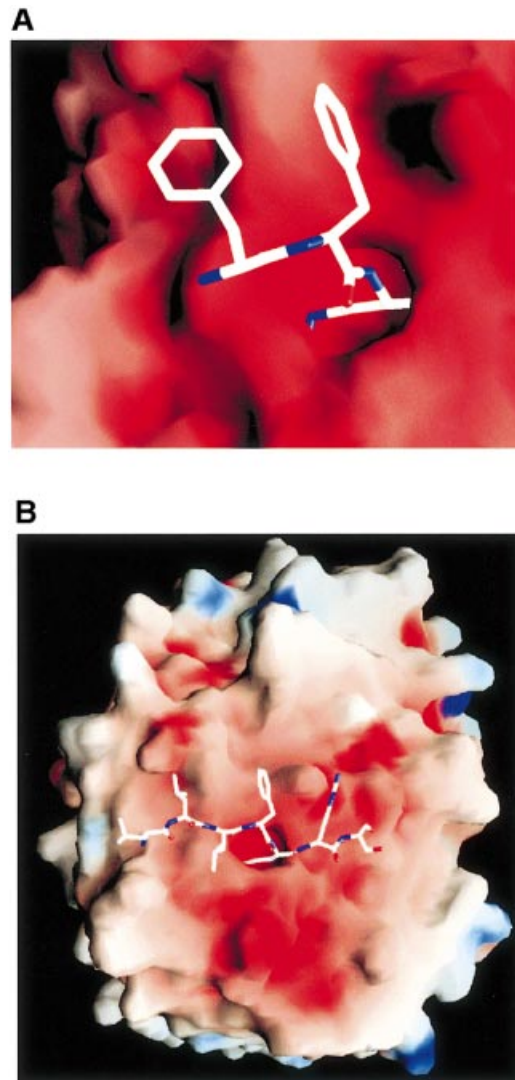


Fig. 4. View toward the solid surface of the RgpB active-site (standard orientation). The electrostatic surface potentials are contoured from –15 (intense red) to 15 $k_B T/e$ (intense blue). Figures were prepared with GRASP (Nicholls *et al.*, 1991). (A) Close-up view toward the active centre of the inhibited RgpB, with the D-Phe–L-Phe–L-Arg methylene moiety (stick model) covalently bound. The Arg side chain is partially buried in the S1 pocket. (B) View toward the active-site of RgpB, with the modelled all-L-heptapeptide non-covalently attached (stick model).

domain, spheres of high electron density in combination with short distances to the coordinating ligands and high coordination numbers indicate the presence of further bound metal ions, which have been interpreted as calcium and zinc ions (FFRCMK only). One calcium is particularly noteworthy in that it is placed just below the S1 specificity pocket (see Figure 1), clamping helices h5 and h9 together and in this way presumably stabilizing this specificity-determining binding site. It is very likely that removal of this calcium ion causes a shift or disordering of Asp163, in this way affecting hydrolytic activity.

Active-site, substrate binding subsites and catalytic mechanism

Figures 3 and 4 allow views into and toward the active centre of RgpB. The active-site and its immediate environment placed on the almost flat 'masticating surface' of the crown is demarcated by strand His211–Glu214 (Figure 3, bottom right), the (intervening) rising segment Ala243–Val245 (right, centre), the (parallel aligned) long s9–h9 connecting segment (top) and the perpendicular running twisted s5–h5 (left, from Glu152 to Asp163). In the centre resides the exposed Cys244, which in the MAD phased map immediately attracted attention due to the neighbourhood to the highest of the three mercury-related peaks accounting for the HgBA molecules. From the bottom strand ejects the His211 imidazole side chain, with its N δ 1 atom in all RgpB structures placed ~5.5 Å away from the Cys244 S γ .

Toward the left of this Cys244 the deep S1 pocket opens, which has the Val242 and Thr209 C γ atoms as a base and is covered by the indole moiety of Trp284 acting as a lid, and at its bottom harbours the Asp163 carboxylate group as a charge anchor. On its upper and lower side, this pocket is lined by the side chain of Met288 and by segment Gly210–His211 and the His211 side chain, respectively (Figure 3). The amidinophenyl part of the HgBA molecule bound to Cys244 extends into this pocket, with its amidino moiety shifted against the Asp163 carboxylate group making only one NH–O hydrogen bond. From the left side, the extended side chain of Glu152 may reach for this His211, so that it could form a short water-exposed hydrogen bond with His211 N ϵ 2 if fully extended. We note, however, that in the crystal structures NATI and pCMBa, the side chain of Glu152 points away from His211. In the FFRCMK structure, Glu152 is connected to His211 N ϵ 2 via a metal ion (probably Zn $^{2+}$ from the crystallization buffer).

In the FFRCMK structure, the peptidic inhibitor is bound to Cys244 forming a covalent bond with Cys244 S γ through its methylene group (Figure 3). The ketone group oxygen forms favourable hydrogen bonds with Gly212 N and Cys244 N, which, together, function as an oxyanion hole. The inhibitor (I) D-PheII–L-Phe2I–L-Arg3I moiety juxtaposes RgpB segment Gln282–Trp284 in a nearly extended manner, forming a two-stranded twisted β -pleated sheet, with a favourable hydrogen bond between Gln282 O and P1-Arg3I NH (3.0 Å), and a longer one between Trp284 NH and P3-PheII O. The Arg3I side chain inserts into the S1 pocket, with its terminal guanidyl group making a frontal 2N–2O salt bridge with the carboxylate group of Asp163, and the N ϵ -H and the two N ζ nitrogens directed toward the Tyr283 carbonyl and the

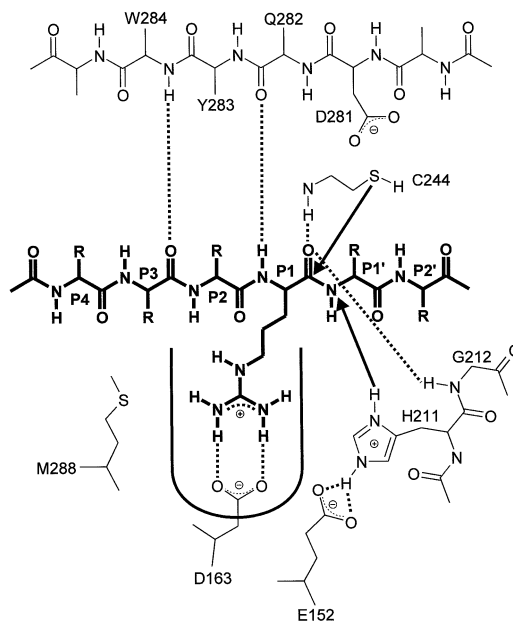


Fig. 5. Schematic drawing of the probable peptide substrate–active centre interaction of RgpB deduced from FFRCMK. The view is in the standard orientation, so that the modelled substrate (thick connections) runs from left to right. Probable intermolecular hydrogen bonds are shown by dashed lines, while the routes of attack of the Cys244 S γ on the Arg-P1 carbonyl and transfer of the His211 N δ hydrogen toward the leaving group are indicated by arrows.

carbonyl groups of Trp284 and Gly210, respectively. This narrow S1 slot, covered by a hydrophobic lid and bordered by in-plane hydrogen bond acceptors, certainly accounts for the strong preference of RgpB for Arg residues at P1/S1. One of both guanidinium N ζ nitrogens is also close enough to Met288 S δ to make a hydrogen bond interaction with one of its lone pair orbitals. In Kgp, according to the sequence alignment of Pavloff *et al.* (1997), this Met288 seems to be replaced by a Tyr residue, whose phenol group might be rotated into the S1 pocket presenting an extra hydrogen bond donor to the ammonium group of an inserted P1-Lys. Thus, residue 288 (together with the basement residue 242) may be a key element in discriminating between the specificities of Rgp and Kgp. The L-Phe2I and D-Phe2I side chains of the inhibitor extend out from the surface of the molecule, with their benzyl groups making edge-on-face contacts with one another and with the phenol group of the adjacent Tyr283 (Figure 3).

Except for the entrance 'hole' to the S1 pocket, the molecular surface around the active-site is relatively flat and characterized by a negative electrostatic potential (Figure 4). Modelling studies show that a productively bound contiguous polypeptide substrate consisting of L-amino acids could bind in an FFRCMK-similar overall extended manner stretching from the N-terminal part of helix h9 (Ala285 N, Trp284 N) along segment Asp281–Tyr283 to strand His211–Glu214, and clamp between P5–S5 and P3'–S3' by up to seven inter-main-chain hydrogen bonds (with some of them shown in Figure 5). The P4 and P2 side chains could nestle toward (hydrophobic/acidic) surface depressions, while the P3 side chain could reside on top of the Trp284 indole lid and the P1' (or maybe P2') side chains making surface contacts with

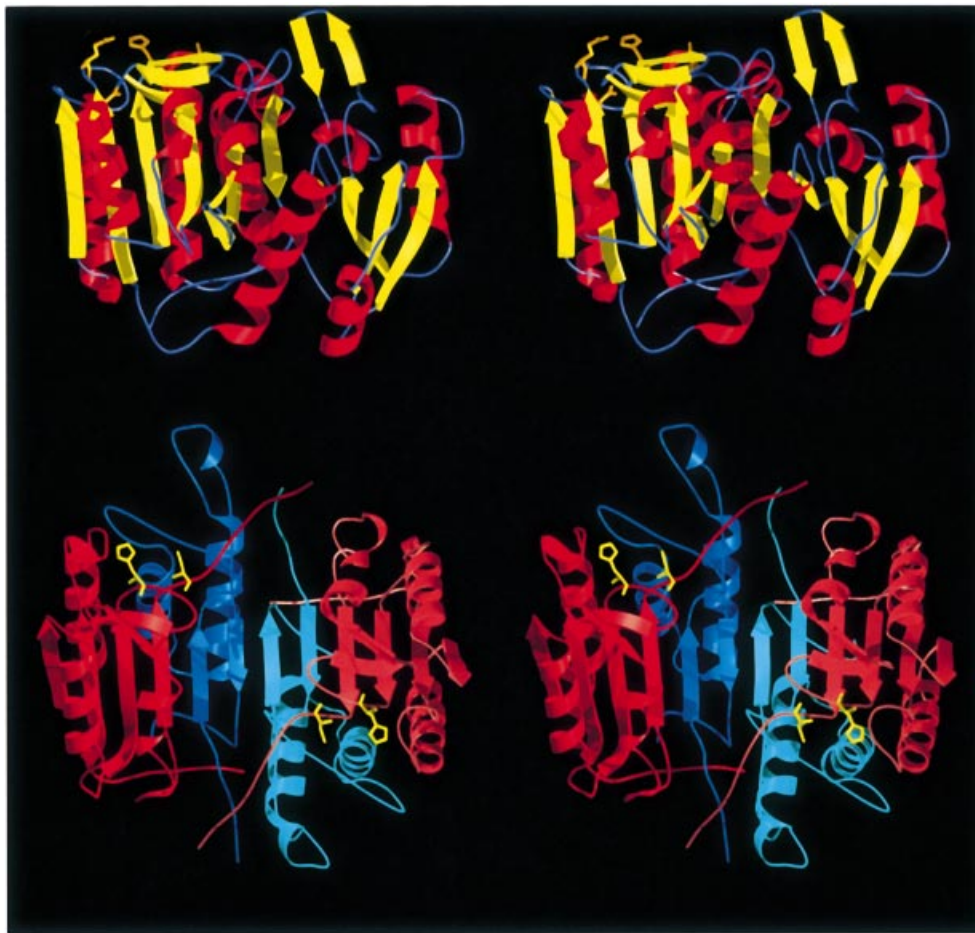


Fig. 6. Comparison of the RgpB catalytic domain (top, in stereo, red, yellow and blue) and the caspase-3 tetramer (bottom, in stereo, with both p17 and p12 peptides given in red and blue; Protein Data Bank code 1PAU; Rotonda *et al.*, 1996). The catalytic residues are shown as orange and yellow stick models. The caspase is presented in such an orientation that its left-side p17–p12 heterodimer half superimposes with the ‘active’ RgpB B-subdomain. Both RgpB subdomains are related to one another by a vertical axis with an $\sim 160^\circ$ rotational and an ~ 15 Å translational component arranged within the plane, while the exact 2-fold rotation axis of caspase-3 relating both heterodimers stands vertically on the plane. The figure was prepared with BOBSCRIPT (Kraulis, 1991) and rendered with RASTER3D (Merrit and Murphy, 1994).

the Asp281 carboxylate. Thus, RgpB would seem to bind its peptide substrates primarily via main-chain interactions, with the exact cleavage site determined, however, by the quite selective P1–Arg–S1 interaction assisted by P4–S4, P2–S2 and P1’–S1’. It should be noted that the S2 pocket residue 283 of RgpA and HRgpA (Swiss-Prot entry code Q45168; Okamoto *et al.*, 1995) is Ser, while in RgpB (Swiss-Prot entry code P95493; Potempa *et al.*, 1998) Tyr occupies this position.

In this way, the carbonyl group of the scissile Arg–Xaa peptide bond of a bound substrate is presented in a rigid and stereochemically favourable manner to Cys244 for nucleophilic attack by Sy (Figure 5). Assisted by the polarization of the P1 carbonyl in the oxyanion hole, Cys244 Sy could bind to the carbonyl carbon of the Arg–Xaa scissile peptide bond toward its *Re* face, under approach of the tetrahedral intermediate state. In this reaction, the attacking Sy lone pair orbital might be oriented toward the carbonyl by hydrogen bonding from the Gln282 N-H. The thiolate anion would certainly be a better nucleophile than the uncharged thiol group in forming the tetrahedral intermediate. However, at neutral pH the active centre exhibits quite a negative electrostatic

surface potential (Figure 4), and the Cys244 side chain is placed adjacent to the exposed carboxylate group of Asp281, so that its thiol group presumably has a normal if not high pK. A putative acidic protonated sulfur of the hemimercaptal part created upon nucleophilic attack by the thiol could easily transfer its proton to the spatially adjacent carboxylate group of Asp281 or to bulk water.

Simultaneously, the imidazole group of His211 positioned on the opposite side of the scissile bond could, possibly, in a concerted move together with the Glu152 carboxylate group, turn toward the pyramidalizing Xaa leaving group nitrogen. The strongly negative electrostatic surface potential together with the properly placed Glu152 would probably stabilize this His211 imidazole in its protonated form, enabling it to donate a proton to the leaving group nitrogen, thereby promoting the C–N break in the bound substrate and the release of the C-terminal fragment (see Figure 5). The thiol ester remaining after release of the C-terminal fragment could be hydrolysed by a water molecule attacking the ester carbonyl and cleaving the ester bond, leading to the release of the N-terminal portion of the substrate under simultaneous transfer of both protons to the His211 imidazole and the Cys244 Sy, respectively.

Comparison with caspases

A topological search of the RgpB structure performed with TOP3D (Lu, 1996) against all protein structures currently in the PDB revealed caspase-1 (Walker *et al.*, 1994; Wilson *et al.*, 1994) and caspase-3 (Rotonda *et al.*, 1996; Mittl *et al.*, 1997) as the only topological homologues (Figure 6). A closer look showed that the A- as well as the B-subdomain each contain a caspase folding motif corresponding to the p20–p10 heterodimer of caspase-1 and the p17–p12 heterodimer of caspase-3. After optimal superposition of gingipain's B-subdomain over the heterodimer of caspase-1 (with ~120 equivalent C α atoms deviating by ~1.5 Å, 17 of which carry identical amino acids; see Figures 2 and 6), strands and helices s5, h5, s6, h6, s7, h7 and s8 of RgpB superimpose well with the first four strands and the first three helices of the caspase-1 p20 chain, with conserved strand order and direction. Only RgpB helix h8 lacks any equivalent in the caspases, but presumably corresponds to the connecting segment of pro-caspases. RgpB strands and helices s9, h9, h10 and s10 are arranged like the last two strands and the two intervening helices of the caspase-1 p10 subunit, with helix h11 replaced by a shorter loop. The topological equivalency of the RgpB A-subdomain is less striking; the first four strands and three helices superimpose well with the p20 subunit of caspase-1, whereas only the first helix of the p10 chain has a corresponding (h4) helix in RgpB.

These two superpositions reflect an internal quasi-symmetry within RgpB, with its A-subdomain showing a high degree of topological similarity to the N-terminal region of the B-subdomain. In contrast to the exact 2-fold axis relating the two heterodimeric halves of caspase-1 and caspase-3, however, both RgpB subdomains are related by an axis running almost perpendicular to the caspase-1 and caspase-3 2-fold, not allowing a simultaneous fit to both caspase heterodimers (Figure 6). If this strong topological similarity reflects an evolutionary relationship between RgpBs and caspases, the RgpB structure would suggest that the caspase heterodimer usually selected (see Figure 6) is indeed generated from one single pro-caspase chain. According to the superposition of the RgpB B-subdomain on the caspase-1 heterodimer (Figures 2B and 6), both the catalytic Cys (244 in RgpB and 285 in caspase-1) and His (211 and 237, respectively) residues occupy identical sites and exhibit very similar conformations. Similarly shaped S1 pockets are carved into the molecular surface, with a frontally opposing acidic residue (RgpB, Asp163) and two embracing basic residues (Arg179 and Arg341 in caspase-1) at the bottom, however. The caspases do not possess a third acidic catalytic residue but a nearby carbonyl (of Pro177 in caspase-1 and of Thr62 in caspase-3), while in RgpB the putative third residue is Glu152. Furthermore, their active-site is placed in a much more pronounced cleft in comparison to RgpB. In contrast to the negative electrostatic surface potential around the active-site of RgpB, the cleft and the surface environment of caspases exhibit a positive potential. Polypeptide substrates seem to interact with the non-primed subsites of caspases with a similar geometry, but with a tighter insertion of the P4 residue into a more pronounced S4 cleft.

Immunoglobulin superfamily domain

At Thr352, the RgpB chain enters the IgSF domain (Figure 1) displaying the shape of an elongated seven-stranded β -barrel (Figures 1 and 2). The strands, referred to according to the IgSF standard nomenclature (see Chothia and Jones, 1997), are connected and arranged as in IgG domains, with the connecting loops being shorter, however. An exception is the sC–sC' β -turn, which projects out of the barrel. Strand sA, after a short crossing over sG, first runs antiparallel to sB, before kinking over and aligning again (as sA') to sB (Figure 2A). The strand following sC primarily makes a β -hairpin with sC itself and is, therefore, best identified as sC'; in its terminal end, however, this strand forms a β -hairpin loop with sE, and thus could be denoted sD (Figure 2A). The IgSF domain of RgpB, therefore, appears to belong to the I set of IgSFs, as is characteristic for bacterial members (see Chothia and Jones, 1997).

The IgSF domain does not contain the characteristic intradomain disulfide bridge usually clamping strands sB and sF (and concomitantly both IgSF sheets) together. However, Cys373 of RgpB occupies an equivalent position in sB, with its unpaired side chain extending into the hydrophobic core where it opposes Val417 of sF. On the right-side domain surface (Figure 1), a depression formed by loops sC–sC' and sE–sF carries a cluster of acidic residues (384-Asp–Asp–Gly–Asp–387 and 405-Glu–Ser–Ile–Ala–Asp–Glu–410, respectively). The sC–sC' loop, above this site, and the sF–sG connection represent the two contact sites through which the IgSF domain props itself against the catalytic domain. The former interactions toward the connecting peptide are quite polar, while the latter, against the h1–s2 loop, are made exclusively via van der Waals contacts. Between these contacts, both domains embrace a water-filled cavity, which is primarily lined by hydrophobic side chains and exits to the back (Figure 1).

Conclusions

The similarity of the RgpB catalytic domain with caspases is certainly not just accidental. The identical arrangement of the secondary structure elements together with an almost identical active-centre topology are a quite significant indicator of a close evolutionary relationship, with the RgpB structure suggesting that the usually selected (red–dark blue, Figure 6) caspase heterodimer is indeed generated from a single pro-caspase chain (see Thornberry and Lazebnik, 1998). A common catalytic ancestor might have had half the size of RgpB catalytic domain, corresponding to a p20–p10 caspase-1 heterodimer; duplication of the ancestral gene, under preservation of only one catalytic centre, might then have led to the present quasi-symmetrical Rgp domain. Both proteinases share identically positioned active Cys and His residues, an equivalent oxyanion hole, similar non-primed and primed recognition subsites and a similarly placed and shaped S1 pocket, serving as the main recognition determinant. In contrast to the Asp-adapted pocket of caspases, however, the S1 pocket of RgpB is optimized to accommodate Arg side chains preferentially. The rigid fixation of bound Arg peptides not only explains the strict specificity of RgpB for P1-Arg-containing peptides, but also accounts for the

high cleavage efficiency of RgpB toward peptide substrates of all lengths.

The catalytic mechanism of RgpB as outlined above (see Figure 5) should be similar to that of the caspases, i.e. with a thiol(ate) in the acylation step making an oxyanion hole-assisted nucleophilic attack on the scissile peptide carbonyl of the peptide substrate, and the spatially separated imidazolium acting as a general acid to donate a proton to the leaving group nitrogen. In stark contrast to the caspases, however, the quite negative surface potential and the presence of residue Glu152 at the RgpB active-site favour stabilization of the imidazolium cation, while disfavoring formation of a thiolate anion. The much higher pH optimum of RgpB (Potempa *et al.*, 1998) compared with caspases (Stennicke and Salvesen, 1997) might reflect the presumed higher pK of the reactive Cys in RgpB.

Under reducing conditions, RgpB seems to be a powerful proteinase, which due to a quite open binding site and strong fixation of the P1-Arg residue is able to hydrolyse a multitude of Arg-Xaa bonds in proteins or peptides. With respect to its cleavage preference for Arg-Xaa, further assisted by the negatively charged S1 and active-site environment, RgpB resembles trypsin, and like many trypsin-like serine proteinases it seems to be capable of cleaving/activating cascade proteinases. However, in contrast to these serine proteinases, RgpB needs an anaerobic environment to be active, as probably exists in the periodontal pockets of infected patients.

The function of the IgSF domain for RgpB is less obvious. In the high-molecular-weight Rgps, this IgSF domain could represent the attachment site for anchoring the haemagglutinin domain. Like many other IgSF domains involved in cell-cell interaction, receptor docking and general antigen recognition and binding (for a review, see Chothia and Jones, 1997), in RgpB this IgSF domain could help bind to protein substrates or dock to endogenous proteins, other bacteria or host cell surfaces. Adhesion and binding of the gingipains seem rather to be mediated via the haemagglutinin domain, which in the high-molecular-weight forms is bound in addition. This is emphasized by the considerably higher activity of HRgpA than RgpB with respect to activation of factor X, mainly effected by a decreased K_m value (Imamura *et al.*, 1997).

Porphyromonas gingivalis has been shown to be a primary pathogen in periodontal disease. Thus, all factors impairing the colonization and emergence of this bacterium would be of fundamental importance to control periodontitis (Holt *et al.*, 1988). The gingipains, in particular the Rgps, are involved in a number of intrinsic and extrinsic functions that are associated with the virulence of *P.gingivalis* and assumed to be essential for its survival. Through their effects on blood clotting and platelet aggregation, these proteinases might be implicated as the biochemical link in the correlation between periodontitis and coronary diseases (Page, 1998). The three-dimensional structure of RgpB now offers the unique possibility to design and elaborate selective potent inhibitors against RgpB and the related proteinases in a rational way. Drugs developed from such inhibitors might be helpful tools to prevent or treat periodontal diseases and to investigate the role of these enzymes further.

In support of our postulations, we have become aware

of a publication (Chen *et al.*, 1998) proposing some relationships within a clan called CD comprising, besides the gingipains (C25), the cysteine proteinase families of the legumains (C13), clostripains (C11) and caspases (C14), mainly on the basis of the sequential His-Gly-spacer-Ala-Cys motif. Gingipain structural analysis now not only confirms some equivalency with caspases, but indicates considerable similarities regarding polypeptide fold and active-site geometry. Furthermore, it suggests closer structural similarities of gingipain with legumains and clostripains as well.

Materials and methods

Enzyme purification and crystallization

RgpB was purified from *P.gingivalis* HG66 as described previously (Potempa *et al.*, 1998). For crystallization, the major RgpB form (pI of 5.1) was selected and activated by incubation with 10 mM L-cysteine in 0.1 M HEPES, 2.5 mM CaCl₂ pH 8.0 at 37°C for 30 min. For alkylation, part of the activated RgpB was treated with the fast-acting irreversible inhibitor FFRCKM (from Bachem; see Potempa *et al.*, 1997). The sample was then dialysed against 3 mM MOPS, 0.02% Na₂S₂O₃ pH 7.2.

Crystallization experiments were carried out with Linbro plates by the vapour diffusion method. For preparation of the monoclinic crystals NATI and pCMBA, drops of 1.5 µl of an 8 mg/ml protein solution and 1.5 µl of reservoir solution (3.6 M 1,6-hexanediol, 300 mM MgCl₂, 100 mM Tris-HCl pH 8.6) were mixed and equilibrated at 6°C for 1 year and then for 6 months at 21°C. For faster crystal growth, microseeding was applied under slightly optimized conditions (3.4 M 1,6-hexanediol, 200 mM MgCl₂, 100 mM Tris-HCl pH 8.5) at 21°C, yielding crystals of up to 0.5 × 0.2 × 0.02 mm within several weeks. These monoclinic crystals (Table I) diffract to beyond 1.5 Å and contain one molecule per asymmetric unit ($V_M = 2.15 \text{ \AA}^3/\text{Da}$), corresponding to a solvent content (v/v) of 43%. The pCMBA derivative was obtained by soaking overnight in a solution containing 3 mM self-made *p*-chloromercuric benzamidine and 5 mM cysteine. Before harvesting, the crystal was washed for 3 min in the precipitant solution to remove unbound mercury atoms.

The RgpB-chloromethylketone complex was crystallized as described previously (Banbula *et al.*, 1998).

Data collection and processing

A NATI dataset to 2.16 Å resolution was measured in-house using a 300 mm MAR-Research image plate detector mounted on a Rigaku RU 200 rotating anode X-ray generator with graphite monochromatized CuK α radiation. MAD measurements were performed on pCMBA crystals at the BW6 beam line at DESY in Hamburg, Germany. From a crystal flash frozen in liquid nitrogen, diffraction data to 1.49 Å resolution were collected at cryo temperatures using a MAR-Research CCD detector. MAD data were measured at the three wavelengths λ_1 (remote, $f' = -8.7 \text{ e}$, $f'' = 4.3 \text{ e}$), λ_2 (peak, $f' = -12.8 \text{ e}$, $f'' = 10.0 \text{ e}$) and λ_3 (edge, $f' = -15.1 \text{ e}$, $f'' = 9.5 \text{ e}$) by continuously collecting 260 (remote), 590 (peak) and 259 (edge) frames of 0.7° each (see Table II). Because the signal of the mercury atoms in the small protein crystal used was too weak, the fluorescence scan to determine optimal wavelengths for data collection was carried out using crystalline chloromercuric benzoic acid (Fluka).

X-ray diffraction data of FFRCKM were collected to 2.0 Å resolution from orthorhombic crystals mounted in glass capillaries at 16°C on a 300 mm MAR-Research image plate detector (Banbula *et al.*, 1998). The intensities of all datasets were integrated with MOSFLM (Leslie, 1991), scaled with SCALA [Collaborative Computational Project Number 4 (CCP4), 1994] and converted to amplitudes with TRUNCATE (CCP4, 1994).

The three mercury atoms in pCMBA (bound to Cys244, Cys299 and Cys185) were localized in an anomalous difference Patterson map of the peak wavelength data using CCP4 programs. The refinement of heavy-atom parameters and calculation of MAD phases were carried out with SHARP (de La Fortelle and Bricogne, 1997). The initial MAD phases were improved with SOLOMON (Abrahams and Leslie, 1996), resulting in an interpretable 1.5 Å electron density map. The NATI model was determined by molecular replacement with MOLREP (Vagin and Teplyakov, 1997) using 3 Å data and the pCMBA model as

Table II. MAD phasing statistics

	R_{Cullis}			Phasing power		
	λ_1 remote	λ_2 peak	λ_3 edge	λ_1 remote	λ_2 peak	λ_3 edge
Centric reflections	–	0.52	0.73	–	1.97	1.21
Acentric reflections						
Iso.	–	0.48	0.68	–	2.67	1.55
Ano.	0.93	0.83	0.83	0.31	1.36	1.20

Number of centric reflections, 2400; number of acentric reflections, 61 334.

Mean figure of merit, 0.452 (16.0 – 1.5 Å).

$R_{\text{Cullis}} = \langle \text{phase integrated lack of closure} \rangle / |F_{\text{PH}} - F_{\text{P}}|$

Phasing power = $|F_{\text{H,calc}}| / \langle \text{phase integrated lack of closure} \rangle$

determined by MAD. The initial R -factor and correlation coefficient were 37.6 and 66.9%. The FFRCMK structure was determined by molecular replacement using the refined pCMBA model, 3 Å data and MOLREP. The initial R -factor and correlation were 35.1 and 70.5%, respectively.

Model building and refinement

A first pCMBA model was built on an Evans and Sutherland graphic workstation with FRODO (Jones, 1978) against a 1.5 Å solvent-flattened electron density map calculated with the MAD phases. This and later models were subjected to crystallographic refinement cycles with CNS (Brünger *et al.*, 1998) using the conjugate gradient method with an amplitude-based maximum-likelihood target function and Engh and Huber (1992) geometric restraints, in between times manually corrected upon visual inspection of the display. In early stages, improved phases obtained with SIGMAA (Read, 1986) by a combination of the initial MAD phases with phases calculated from the current partial model and solvent flattening (SOLOMON; Abrahams and Leslie, 1996) were used for structure factor and density calculations. In the first step of the refinement process, each initial model was subjected to simulated annealing. Later, a water model was calculated using ARP (Lanzin and Wilson, 1993), and individual isotropic B -factors were refined using refinement protocols employing an amplitude-based maximum-likelihood target function until convergence was reached.

For the refinement of the orthorhombic FFRCMK model, the appropriately placed pCMBA model, with all residues around the active-site cysteine omitted, served to calculate a simulated annealing omit map. The corresponding 2.0 Å density map allowed tracing of all omitted residues and of the complete inhibitor using usual Arg-methylene parameters and a restrained distance of 1.80 Å for C–S γ . After remodeling and completely new insertion of solvent molecules, this model was refined as carried out for pCMBA. The final refinement statistics are given in Table III.

In the final NATI and pCMBA models, the polypeptide chain (except for Glu66–Gly67) is defined continuously from Tyr1 to Ser435. The C-terminus extends into a large crystal cavity that could accommodate some more residues beyond Ser435 if present. The final FFRCMK model additionally lacks residues 433–435. Besides the protein chains and the inhibitors, the models contain a few ions interpreted as calcium and zinc ions (Table III). The peptide groups preceding Pro49 and Pro188 have *cis* conformation. According to a search performed with PROCHECK (Laskowski *et al.*, 1993), almost all non-glycine residues fall into the allowed or additionally allowed regions of the Ramachandran plot (Table III); however, in all three structures, Val245 (ϕ, ψ pair around 35/13), Ser220 (105/–30) and Lys326 (50/–120) are clearly in the generously allowed region.

The coordinates of FFRCMK have been deposited in the Protein Data Bank under code 1CVR.

Acknowledgements

We thank Drs Hans Bartunik and Gleb P. Bourenkov, DESY, Hamburg, for their excellent help in collecting the MAD data. Financial support by the NIH (grant DE 09761, J.T.), the Polish Committee of Scientific Research (grant 6 P204A 019 11, KBN, Poland, J.P.) and by the SFB469 of the University of Munich, the Biotech and Training-and-Mobility programmes (BIO4-CT98-0418; ERBFMRXCT98-0193) of the

Table III. Refinement statistics

Model	pCMBA	NATI	FFRCMK
Resolution range (Å)	17.0–1.5	17.0–2.16	15.0–2.0
Reflections in working set	110 620	16 891	25 146
Reflections in test set	5610	1848	1865
R_{cryst} (%) ^a	16.6	17.4	16.3
R_{free} (%) ^a	19.1	24.0	20.7
Protein atoms (non-H)	3372	3357	3346
Ligand atoms (non-H)	21	–	42
Water molecules	561	237	322
Ion atoms	3 Ca	3 Ca	3 Ca, 4 Zn
Average B -factor (Å ²)	16.1	20.2	22.3
R.m.s. ΔB (Å ²)	3.0	3.0	1.8
Deviations from ideality			
R.m.s. bonds (Å)	0.014	0.010	0.009
R.m.s. angles (°)	1.79	1.64	1.55
Model quality due to Ramachandran analysis			
Allowed (%)	90.3	89.8	89.4
Additionally allowed (%)	8.9	9.2	9.8
Generously allowed (%)	0.8	1.0	0.8

$$^a R_{\text{cryst,free}} = \Sigma (|F_{\text{obs}} - F_{\text{calc}}|) / \Sigma |F_{\text{obs}}|$$

European Union and the HFSP programme (RG-203/98) (W.B.) is kindly acknowledged.

References

- Abrahams, J.P. and Leslie, A.G.W. (1996) Methods used in the structure determination of bovine mitochondrial F₁ ATPase. *Acta Crystallogr. D*, **52**, 30–42.
- Banbula, A., Potempa, J., Travis, J., Bode, W. and Medrano, F.-J. (1998) Crystallization and preliminary X-ray diffraction analysis of gingipain R2 from *Porphyromonas gingivalis* in complex with H-D-Phe-Phe-Arg-chloromethylketone. *Protein Sci.*, **7**, 1259–1261.
- Barton, G.J. (1993) ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.*, **6**, 37–40.
- Brünger, A.T. *et al.* (1998) Crystallography & NMR System—a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.
- Calkins, C.C., Platt, K., Potempa, J. and Travis, J. (1998) Inactivation of tumor necrosis factor- α by proteinases (gingipains) from the periodontal pathogen, *Porphyromonas gingivalis*. *J. Biol. Chem.*, **273**, 6611–6614.
- Cerretti, D.P. *et al.* (1992) Molecular cloning of the interleukin-1 β converting enzyme. *Science*, **256**, 97–100.
- Chen, J.-M., Rawlings, N.D., Stevens, R.A.E. and Barrett, A.J. (1998) Identification of the active site of legumain links it to caspases, clostripain and gingipains in a new clan of cysteine endopeptidases. *FEBS Lett.*, **441**, 361–365.
- Chen, Z., Potempa, J., Polanowski, A., Wikstrom, M. and Travis, J. (1992) Purification and characterization of a 50-kDa cysteine proteinase

- (gingipain) from *Porphyromonas gingivalis*. *J. Biol. Chem.*, **267**, 18896–18901.
- Chothia, C. and Jones, E.Y. (1997) The molecular structure of cell adhesion molecules. *Annu. Rev. Biochem.*, **66**, 823–862.
- Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.
- Cutler, C.W., Kalmar, J.R. and Genco, C.A. (1995) Pathogenic strategies of the oral anaerobe, *Porphyromonas gingivalis*. *Trends Microbiol.*, **3**, 45–51.
- de La Fortelle, E. and Bricogne, G. (1997) Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.*, **276**, 472–494.
- DiScripio, R.G., Daffern, P.J., Kawahara, M., Pike, R., Potempa, J., Travis, J. and Hugli, T.E. (1996) Cleavage of human complement component C5 by cysteine proteinases from *Porphyromonas (Bacteroides) gingivalis*. Prior oxidation of C5 augments proteinase digestion of C5. *Immunology*, **87**, 660–667.
- Engh, R.A. and Huber, R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A*, **47**, 392–400.
- Fujimura, S., Hirai, K., Shibata, Y., Nakayama, K. and Nakamura, T. (1998) Comparative properties of envelope-associated arginine-gingipains and lysine-gingipain of *Porphyromonas gingivalis*. *FEMS Microbiol. Lett.*, **163**, 173–179.
- Genco, C.A., Odusanya, B.M., Mikolajczyk-Pawlinska, J., Potempa, J. and Travis, J. (1998) A peptide domain on gingipain R which confers immunity against *Porphyromonas gingivalis* infection in mice. *Infect. Immun.*, **66**, 4108–4114.
- Holt, S.C., Ebersole, J., Felton, J., Brunsvold, M. and Kornman, K.S. (1988) Implantation of *Bacteroides gingivalis* in nonhuman primates initiates progression of periodontitis. *Science*, **239**, 55–57.
- Imamura, T., Pike, R.N., Potempa, J. and Travis, J. (1994) Pathogenesis of periodontitis: a major arginine-specific cysteine proteinase from *Porphyromonas gingivalis* induces vascular permeability enhancement through activation of the kallikrein/kinin pathway. *J. Clin. Invest.*, **94**, 361–367.
- Imamura, T., Potempa, J., Tanase, S. and Travis, J. (1997) Activation of blood coagulation factor X by arginine-specific cysteine proteinases (gingipain-Rs) from *Porphyromonas gingivalis*. *J. Biol. Chem.*, **272**, 16062–16067.
- Jones, T.A. (1978) A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.*, **11**, 268–272.
- Kadowaki, T., Nakayama, K., Yoshimura, F., Okamoto, K., Abe, N. and Yamamoto, K. (1998) Arg-gingipain acts as a major processing enzyme for various cell surface proteins in *Porphyromonas gingivalis*. *J. Biol. Chem.*, **273**, 29072–29076.
- Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.
- Kuramitsu, H.K. (1998) Proteases of *Porphyromonas gingivalis*: what don't they do? *Oral Microbiol. Immunol.*, **13**, 263–270.
- Lamont, R.J. and Jenkinson, H.F. (1998) Life below the gum line: pathogenic mechanisms of *Porphyromonas gingivalis*. *Microbiol. Mol. Biol. Rev.*, **62**, 1244–1263.
- Lamzin, V.S. and Wilson, K.S. (1993) Automated refinement of protein models. *Acta Crystallogr. D*, **49**, 129–147.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.
- Leslie, A.G.W. (1991) Molecular data processing. In Moras, D., Podjarny, A.D. and Thierry, J.C. (eds), *Crystallographic Computing 5*. Oxford University Press, Oxford, UK, pp. 50–61.
- Lu, G. (1996) A WWW service system for an automatic comparison of protein structures. *Protein Data Bank Q. Newsl.*, **78**, 10–11.
- Merrit, E.A. and Murphy, M.E.P. (1994) Raster3D version 2.0—a program for photorealistic molecular graphics. *Acta Crystallogr. D*, **50**, 869–873.
- Mikolajczyk-Pawlinska, J., Kordula, T., Pavloff, N., Pemberton, P.A., Kiefer, M.C., Travis, J. and Potempa, J. (1998) Genetic variations of *Porphyromonas gingivalis* genes encoding gingipains, cysteine proteinases with arginine or lysine specificity. *Biol. Chem.*, **379**, 205–211.
- Mittl, P.R.E., DiMarco, S., Krebs, J.F., Bai, X., Karanewsky, D.S., Priestle, J.P., Tomaselli, K.J. and Gruetter, M.G. (1997) Structure of recombinant human CPP32 in complex with the tetrapeptide Acetyl-Asp-Val-Ala-Asp fluoromethyl ketone. *J. Biol. Chem.*, **272**, 6539–6547.
- Nakayama, K. (1997) Domain-specific rearrangement between two arg-gingipain-encoding genes in *Porphyromonas gingivalis*: possible involvement of nonreciprocal recombination. *Microbiol. Immunol.*, **41**, 185–196.
- Nakayama, K., Kadowaki, T., Okamoto, K. and Yamamoto, K. (1995) Construction and characterization of arginine-specific cysteine proteinase (Arg-gingipain)-deficient mutants of *Porphyromonas gingivalis*. Evidence for significant contribution of Arg-gingipain to virulence. *J. Biol. Chem.*, **270**, 23619–23626.
- Nicholls, A., Sharp, K. and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.*, **11**, 281–296.
- Nishikata, M. and Yoshimura, F. (1995) Active site structure of a hemagglutinating protease from *Porphyromonas gingivalis*: similarity to clostripain. *Biochem. Mol. Biol. Int.*, **37**, 547–553.
- Okamoto, K., Misumi, Y., Kadowaki, T., Yoneda, M., Yamamoto, K. and Ikehara, Y. (1995) Structural characterization of argingipain, a novel arginine-specific cysteine proteinase as a major periodontal pathogenic factor from *Porphyromonas gingivalis*. *Arch. Biochem. Biophys.*, **316**, 917–925.
- Onoe, T., Hoover, C.I., Nakayama, K., Ideka, T., Nakamura, H. and Yoshimura, F. (1996) Identification of *Porphyromonas gingivalis* pre-fimbriin possessing a long leader peptide: possible involvement of trypsin-like protease in fimbriin maturation. *Microbiol. Pathogen.*, **19**, 351–364.
- Page, R.C. (1998) The pathobiology of periodontal diseases may affect systemic diseases: inversion of a paradigm. *Ann. Periodontol.*, **3**, 108–120.
- Pavloff, N., Potempa, J., Pike, R.N., Prochazka, V., Kiefer, M.C., Travis, J. and Barr, P.J. (1995) Molecular cloning and structural characterization of the Arg-gingipain proteinase of *Porphyromonas gingivalis*. *J. Biol. Chem.*, **270**, 1007–1010.
- Pavloff, N., Pemberton, P.A., Potempa, J., Chen, W.C.A., Pike, R.N., Prochazka, V., Kiefer, M.C., Travis, J. and Barr, P.J. (1997) Molecular cloning and characterization of *Porphyromonas gingivalis* Lys-gingipain. A new member of an emerging family of pathogenic bacterial cysteine proteinases. *J. Biol. Chem.*, **272**, 1595–1600.
- Potempa, J. and Travis, J. (1996) *Porphyromonas gingivalis* proteinases in periodontitis, a review. *Acta Biochim. Pol.*, **43**, 455–466.
- Potempa, J., Pike, R. and Travis, J. (1997) Titration and mapping of the active site of cysteine proteinases from *Porphyromonas gingivalis* (gingipains) using peptidyl chloromethanes. *Biol. Chem.*, **378**, 223–230.
- Potempa, J., Mikolajczyk-Pawlinska, J., Brassell, D., Nelson, D., Thøgersen, I.B., Enghild, J.J. and Travis, J. (1998) Comparative properties of two cysteine proteinases (gingipain Rs), the products of two related but individual genes of *Porphyromonas gingivalis*. *J. Biol. Chem.*, **273**, 21648–21657.
- POV-Ray Program, Version 3.02 (1997) URL: <http://www.povray.org>
- Rangarajan, M., Aduse-Opoku, J., Slaney, J.M., Young, K.A. and Curtis, M.A. (1997) The prpR1 and prR2 arginine-specific protease genes of *Porphyromonas gingivalis* W50 produce five biochemically distinct enzymes. *Mol. Microbiol.*, **23**, 955–965.
- Rano, T.A., Timkey, T., Peterson, E.P., Rotonda, J., Nicholson, D.W., Becker, J.W., Chapman, K.T. and Thornberry, N.A. (1997) A combinatorial approach for determining protease specificities: application to interleukin-1 β converting enzyme (ICE). *Chem. Biol.*, **4**, 149–155.
- Read, J.M. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A*, **42**, 140–149.
- Rotonda, J. et al. (1996) The three-dimensional structure of apopain/CPP32, a key mediator of apoptosis. *Nature Struct. Biol.*, **3**, 619–625.
- Stennicke, H.R. and Salvesen, G.S. (1997) Biochemical characteristics of caspases-3, -6, -7 and -8. *J. Biol. Chem.*, **272**, 25719–25723.
- Thornberry, N.A. and Lazebnik, Y. (1998) Caspases: enemies within. *Science*, **281**, 1312–1316.
- Thornberry, N.A. et al. (1992) A novel heterodimeric cysteine protease is required for interleukin-1 β processing in monocytes. *Nature*, **356**, 768–774.
- Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.*, **30**, 1022–1025.
- Walker, N.P. et al. (1994) Crystal structure of the cysteine protease interleukin-1 β -converting enzyme: a (p20/p10) $_2$ homodimer. *Cell*, **78**, 343–352.
- Wilson, K.P. et al. (1994) Structure and mechanism of interleukin-1 β converting enzyme. *Nature*, **370**, 270–275.

Received July 13, 1999; revised and accepted August 31, 1999