Critical role of amino acid 23 in mediating activity and specificity of vinckepain-2, a papain-family cysteine protease of rodent malaria parasites

Ajay SINGH*, Bhaskar R. SHENAI*, Youngchool CHOE†, Jiri GUT*, Puran S. SIJWALI*, Charles S. CRAIK† and Philip J. ROSENTHAL*[1]

*Department of Medicine, San Francisco General Hospital, San Francisco, CA 94143, U.S.A., and †Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143, U.S.A.

Cysteine proteases of *Plasmodium falciparum*, known as falcipains, have been identified as haemoglobinases and potential drug targets. As anti-malarial drug discovery requires the analysis of non-primate malaria, genes encoding related cysteine proteases of the rodent malaria parasites *P. vinckei* (vinckepain-2) and *P. berghei* (berghepain-2) were characterized. These genes encoded fairly typical papain-family proteases, but they contained an unusual substitution of Gly^{23} with Ala (papain numbering system). Vinckepain-2 was expressed in *Escherichia coli*, solubilized, refolded and autoprocessed to an active enzyme. The protease shared important features with the falcipains, including an acidic pH optimum, preference for reducing conditions, optimal cleavage of peptide substrates with P2 Leu and ready hydrolysis of haemoglobin. However, key differences between the plasmodial proteases were identified. In particular, vinckepain-2 showed very different kinetics against many substrates and an unusual preference for peptide substrates with P1 Gly. Replacement of Ala^{23} with Gly remarkably altered vinckepain-2, including loss of the P1 Gly substrate preference, markedly increased catalytic activity (k_{cat}/K_m increased approx. 100-fold) and more rapid autohydrolysis. The present study identifies key animal-model parasite targets. It indicates that drug discovery studies must take into account important differences between plasmodial proteases and sheds light on the critical role of amino acid 23 in catalysis by papain-family proteases.

Key words: falcipain, haemoglobin, *Plasmodium*, proteinase.

INTRODUCTION

Malaria is one of the most important infectious diseases in the world. Infection with *Plasmodium falciparum*, the most virulent human malaria parasite, is responsible for hundreds of millions of illnesses and over one million deaths each year [1]. Treatment of malaria has relied on a small number of drugs, and the emergence of drug-resistant parasites poses a major challenge. Thus the identification and characterization of new targets for anti-malarial chemotherapy is an urgent priority.

Proteases play a key role in the life cycles of malaria parasites [2]. In erythrocytic-stage parasites, aspartic [3–5], cysteine [6–8] and metalloproteases [9] all appear to participate in the hydrolysis of haemoglobin, which is required to provide amino acids for parasite protein synthesis. Three papain-family cysteine proteases, known as falcipains, have been identified in *P. falciparum*. Falcipain-2 [7] and falcipain-3 [8] have been recently well characterized. Both enzymes are haemoglobinases that appear to localize to the trophozoite food vacuole, the site of haemoglobin hydrolysis. Cysteine protease inhibitors block haemoglobin hydrolysis and arrest parasite development *in vitro* [6,10] and *in vivo* [11,12], suggesting that plasmodial cysteine proteases are appropriate new chemotherapeutic targets.

Drug discovery is facilitated by well-characterized animal-model systems for studying the *in vivo* efficacy of drugs. As *P. falciparum* cannot be maintained in non-primates, murine models are routinely used for *in vivo* screens of potential anti-malarials. For this reason, we have evaluated falcipain orthologues from rodent malaria parasites. We now report the identification and characterization of vinckepain-2 and berghepain-2, papain-family cysteine proteases from the rodent parasites *P. vinckei* and *P. berghei* respectively. These enzymes share key sequence features with falcipain-2 and -3. However, they contain a substitution of a highly conserved papain-family residue (Gly^{23} is replaced by Ala; papain numbering system). The biochemical properties of vinckepain-2 were similar to those of the falcipains, but the proteases differed significantly in their extended fine specificity and sensitivity to inhibitors. Replacement of Ala^{23} with Gly had a profound impact on the activity and substrate specificity of vinckepain-2, demonstrating the key role of this amino acid in catalysis.

EXPERIMENTAL

Materials

Synthetic peptide substrates were obtained from GlaxoSmithKline and Bachem (Torrance, CA, U.S.A.). Fluoromethyl ketones were obtained from Prototek (Dublin, CA, U.S.A.) and vinyl sulphones from Axys Pharmaceuticals (South San Francisco, CA, U.S.A.). All other biochemical reagents were obtained from Sigma.

Parasites

Frozen stocks of *P. vinckei* and *P. berghei* were used to infect Swiss Webster mice by intraperitoneal injection. Parasites were subsequently passaged in mice by the intraperitoneal injection of 10^5–10^6 infected erythrocytes. Parasitaemias were evaluated on Giemsa-stained smears. Infected blood samples (30–50 %

Abbreviations used: AMC, 7-*N*-4-methylcoumarin; DTT, dithiothreitol; E-64, *trans*-epoxysuccinyl-L-leuciloamido-(4-guanidino)butane; IPTG, isopropyl β-D-thiogalactoside; Ni-NTA, Ni^{2+}-nitrilotriacetate; ORF, open reading frame.

[1] To whom correspondence should be addressed (e-mail rosnthl@itsa.ucsf.edu).

The nucleotide sequences reported in the present study will appear in GenBank® EMBL, DDBJ and GSDB Nucleotide Sequence Databases under accession nos. AY063763 and AY063764.

parasitaemia) were passed through a cell strainer and a Plasmodipur filter (Euro-Diagnostica, Beijerinckweg, The Netherlands) and then centrifuged at 1200 ***g*** for 10 min. To evaluate parasite-soluble extracts, infected erythrocytes were washed with ice-cold PBS, lysed with 0.1 % (w/v) saponin (which lyses erythrocyte, but not parasite, membranes) in ice-cold PBS for 5 min, centrifuged (12000 ***g*** for 10 min at 4 °C) and washed three times with ice-cold PBS; pellets were then processed for isolation of nucleic acids or resuspended in SDS/PAGE sample buffer, boiled for 5 min and centrifuged at 12000 ***g*** for 10 min.

Isolation of DNA and RNA

To prepare genomic DNA, saponin-lysed infected erythrocytes were treated with lysis buffer [100 μg/ml protease-K, 10 mM Tris/HCl (pH 8.0), 100 mM EDTA, 0.5 % (w/v) SDS], followed by phenol extraction and propan-2-ol precipitation. Total RNA was extracted from saponin-lysed infected erythrocytes using Trizol according to the manufacturer's instructions (Life Technologies). The RNA pellet was suspended in sterile water [with 0.1 % (v/v) diethyl pyrocarbonate] and treated with DNase-I. Purity and quantification of nucleic acids were assessed spectrophotometrically. To produce cDNA, RNA samples (free from genomic DNA, as documented by PCR with vinckepain-2- and berghepain-2-specific primers) were reverse-transcribed using the SuperScript preamplification system (Life Technologies) according to the manufacturer's instructions. Briefly, 10–15 μg of total RNA was annealed with the oligo$(dT)_{12–18}$ primer and extended with SuperScript-II reverse transcriptase at 42 °C for 50 min. cDNA samples were treated with RNase-H before PCR amplification.

Southern and Northern blotting

For Southern-blot analysis, genomic DNAs were digested with restriction endonucleases, electrophoresed and transferred to nylon membranes (Amersham Pharmacia Biotech). The regions encoding mature vinckepain-2 (837 bp) and berghepain-2 (834 bp) were amplified using vinckepain-2-specific primers (A3, 5′-AATATGGATCCAAAAAAGCTATTAATTCATTTAGTGACTTG-3′ and E1, 5′-AATAGTCGACTTATTCAATTACAGGAGCAAAAGC-3′) and berghepain-2-specific primers (A5, 5′-AATCAGGATCCTACAAAAGGTATTAATGCATTTAGTGATATG and E2, 5′-TAATAAAGCTTTTATTCAATTATAGGAGCATAACC-3′), purified using the QIAquick GelExtractionKit (Qiagen) and labelled with [α-^{32}P]dATP (Amersham Pharmacia Biotech) using the Random Primers DNALabeling System (Life Technologies). Membranes were prehybridized in 5 × SSPE [1 × SSPE = 0.15 M NaCl/0.01 M NaH_2PO_4 (pH 7.4)/1 mM EDTA]/0.5 % SDS/5 × Denhardt's solution (1 × Denhardt's = 0.02 % Ficoll 400/0.02 % polyvinylpyrrolidone/0.002 % BSA) for 5 h at 55 °C, hybridized overnight at 65 °C in prehybridization buffer containing the specific probe and washed twice at 25 °C for 5 min with 2 × SSC (1 × SSC = 0.15 M NaCl/0.15 M sodium citrate)/0.1 % SDS, once at 55 °C for 15 min with 2 × SSC/0.1 % SDS and once at 60 °C for 10 min with 0.2 × SSC/0.1 % SDS.

For Northern-blot analysis, 10 μg of *P. vinckei* or *P. berghei* total RNA was electrophoresed on 1 % (w/v) agarose–formaldehyde gels and transferred to nylon membranes. The membranes were prehybridized as described for Southern-blot analysis. Hybridizations were performed overnight at 55 °C with the same probes described for Southern-blot analysis, and membranes were washed with 2 × SSC/0.1 % SDS at 25 °C, 1 × SSC/0.1 % SDS at 55 °C and 0.2 × SSC/0.1 % SDS at 55 °C, each for 15 min.

Amplification, cloning and analysis of vinckepain-2 and berghepain-2

A BLAST search of the National Center for Biotechnology Information *Plasmodium* database identified a *P. berghei*-expressed sequence tag with similarity to falcipain-2 (University of Florida Malaria Genome Tag Sequencing Project). The sequence was amplified from *P. berghei* genomic DNA with specific primers (A1, 5′-CCAATATGCTATTAGAAAAAATC-3′ and B, 5′-GTGCTTATATTTTCCCTATCGG-3′). The same primers were used to amplify a homologous fragment from *P. vinckei* genomic DNA. The amplified fragments were purified with the QIAquick Gel Extraction Kit (Qiagen) and cloned into the pCR2.1-TOPO vector using the TA cloning kit (Invitrogen). The recombinant plasmids were purified from positive clones using a QIAprep Spin Mini Prep Kit (Qiagen), and sequences were confirmed for these and the plasmids mentioned below by dideoxy sequencing in both directions.

The upstream region of the vinckepain-2 gene was PCR-amplified from a λ ZAP-II genomic *P. vinckei* library [13] as follows. An aliquot of the library was incubated at 70 °C for 5 min and PCR was performed using one internal vinckepain-2-sequence-specific primer (A2, 5′-CATCCATGTCTATAAGATCTTC-3′) and a λ ZAP-II-specific primer (C, 5′-AATTAACCCTCACTAAAGGG-3′). The amplified fragment was cloned into the pCR2.1-TOPO vector (pTOP-UVP2) and its sequence was confirmed.

The full vinckepain-2 and berghepain-2 open reading frames (ORFs) were amplified from cDNAs of *P. vinckei* and *P. berghei* using primers (D, 5′-CACCATGAGTTATCATTCTAGC-3′ and B) based on the *P. vinckei* genomic sequence. The amplified fragments were cloned in pCR2.1-TOPO, and sequences of recombinant plasmids (pTOP-VP2 and pTOP-BP2) were confirmed from multiple positive clones. Sequence data were analysed at the European Bioinformatics Institute (EMBL) website using the CLUSTAL W method [14].

Expression, purification and refolding of recombinant proteases

A vinckepain-2 (−73 VP2) fragment encoding the entire mature protease and the 73 C-terminal amino acids of the prodomain was amplified from pTOP-VP2 using vinckepain-2-specific primers (A4, 5′-CAATTTGGATCCAGCAATTAGAAGCTGTGAA-3′ and E1). The PCR fragment was digested with *Bam*HI and *Sal*I, gel-purified and ligated into the *E. coli* expression vector pQE-30 (Qiagen) to generate pQ-VP2. The construct was transformed into M15 (pREP4) *E. coli*. DNA was isolated and sequenced from multiple positive clones to confirm insert sequence. A positive clone was grown overnight, treated with isopropyl β-D-thiogalactoside (IPTG) for 4 h, and inclusion bodies were prepared by suspension of cells in 2 M urea/2.5 % (v/v) Triton X-100, sonication and centrifugation at 12000 ***g*** for 20 min. The enzyme was purified from solubilized [in 6 M guanidinium chloride/20 mM Tris/HCl/500 mM NaCl/10 mM imidazole (pH 8.0)] inclusion bodies by Ni^{2+}-nitrilotriacetate (Ni-NTA; Qiagen) chromatography (utilizing a His tag encoded by the vector). Refolding of the purified recombinant protease was optimized by testing over 100 buffer combinations, as described earlier for falcipains [15]. For large-scale refolding, 100 mg of Ni-NTA-purified vinckepain-2 was reduced with dithiothreitol (DTT; 10 mM), diluted 100 times in 2 litres of ice-cold optimized refolding buffer [100 mM Tris/HCl (pH 8.5)/

1 mM EDTA/400 mM L-arginine/HCl/1 mM GSH/0.5 mM GSSG] and incubated at 4 °C for 20 h. The refolded sample was concentrated to 100 ml using a High-Performance Ultrafiltration Cell (Model 2000; Amicon, Danvers, MA, U.S.A.) with a 10 kDa cut-off membrane.

To allow processing to active enzyme, the pH of the refolded sample was adjusted to 5.5 with 3.5 M sodium acetate (pH 2.6) and DTT was added to a final concentration of 5 mM; the precipitated material was removed (0.22 μm filter; Millipore, Bedford, MA, U.S.A.) and the sample was incubated at 37 °C for 2 h. The pH was then readjusted to 6.5 with 1 M Tris/HCl (pH 8.0) and the protein was applied to a Q-Sepharose column (Amersham Pharmacia Biotech) pre-equilibrated with 20 mM Bis-Tris/HCl (pH 6.5) and maintained at 4 °C. The column was washed with 5–10 bed volumes of the same buffer, and the protein was eluted with a 0–0.4 M linear NaCl gradient over 30 min at a flow rate of 1.5 ml/min. Enzyme-containing fractions, monitored by hydrolysis of benzyloxycarbonyl-Leu-Arg-7-*N*-4-methylcoumarin (Z-Leu-Arg-AMC) were pooled and concentrated (10 kDa cut-off Centriprep; Millipore), an equal volume of glycerol was added and the enzyme was stored at −20 °C. For N-terminal sequencing of the mature enzyme, processed and purified vinckepain-2 was electrophoresed, transferred to an Immobilon-P^{SQ} membrane (Millipore), stained with Coomassie Blue, excised and evaluated by Edman sequencing at the Protein and Nucleic Acid Facility, Stanford University Medical Center (Stanford, CA, U.S.A.). Falcipain-2 and -3 were expressed and purified as described previously [7,8].

Construction of mutant vinckepain-2

A mutant vinckepain-2, in which Ala^{23} was replaced by Gly (papain numbering system; vinckepain-2$_{A23G}$; A23G, Ala^{23} → Gly) was generated by the overlap extension method [16,17] in two steps. Firstly, DNA fragments coding for the prodomain (primers D and E3, 5′-GAAAATGCCCAACAT-GA***TC***CGCATTTTTTTTG-3′) and mature protease domain (primers F, 5′-CAAAAAAAATGC***GA***TCATGTTGGGCATT-TC-3′ and B) were amplified from pTOP-VP2 (bases encoding the altered codon are in bold face). Secondly, the fragments were gel-purified, mixed in equimolar ratio and recombined using primers (A4 and E1) for amplification of vinckepain-2$_{A23G}$. The vinckepain-2$_{A23G}$ fragment (−73 mVP2) was cloned into pQE-30, and the mutated sequence of this construct was confirmed. Expression and refolding of vinckepain-2$_{A23G}$ were performed exactly as for the wild-type enzyme.

Production of antibody and immunoblotting

Rats were immunized intraperitoneally with an emulsified polyacrylamide gel slice, containing Ni-NTA-purified recombinant vinckepain-2 in complete (on day 0) or incomplete (days 14, 35 and 68) Freund's adjuvant. Rats were killed and their serum collected on day 82. For immunoblotting, parasite extracts, collected as described above, and refolded vinckepain-2 were separated by SDS/PAGE (10% gel) and transferred to a PVDF membrane. The membrane was blocked with 0.2% I-block (Tropix, Bedford, MA, U.S.A.), washed, incubated with 1:20000 rat anti-vinckepain-2 serum, washed again, incubated with 1:10000 alkaline phosphatase-conjugated goat anti-rat IgG (Jackson Immunoresearch Laboratories, West Grove, PA, U.S.A.) at 37 °C for 1 h and washed again. Reactions were developed using the Western-Star Chemiluminescence Kit (Tropix), as described by the manufacturer.

Haemoglobinase activity assay

To evaluate haemoglobinase activity, 30 μl of the mixture containing 3 μg of native human or rat haemoglobin in 100 mM sodium acetate (pH 5.5)/2 mM GSH and 600 nM vinckepain-2 (or lacking GSH or enzyme, or with added protease inhibitors) was allowed to react at 37 °C for 3 h. Reactions were stopped by the addition of 10 μl of 4 × SDS/PAGE sample buffer, and reaction products were resolved by SDS/PAGE. Spectrophotometric determination of haemoglobinase activity was performed as described previously [18].

Determination of substrate specificity using positional-scanning tetrapeptide libraries

Two synthetic combinatorial libraries were used to determine the substrate specificities of the S1–S4 subsites of vinckepain-2 and vinckepain-2$_{A23G}$ as described previously [19]. To determine P1-specificity, a P1-diverse library consisting of 20 sublibraries was used. In each sublibrary, the P1 position contained one native amino acid, and the P2, P3 and P4 positions were randomized with equimolar mixtures of amino acids (in each case, cysteine was omitted and methionine was replaced by norleucine) for a total of 6859 tetrapeptide substrate sequences per sublibrary. Aliquots of 5×10^{-9} mol from each sublibrary were added to 20 wells of a 96-well Microfluor-1 U-bottom plate (Dynex Technologies, Chantilly, VA, U.S.A.) at a final concentration of 7.3 nM of each compound per well.

To determine P2, P3 and P4 specificity, a complete diverse library was used. In this library, the P2, P3 or P4 position was spatially addressed with 20 amino acids (norleucine was substituted for cysteine), whereas the remaining three positions were randomized. Aliquots of 2.5×10^{-8} mol from each sublibrary were added to 60 wells (8000 compounds/well) of a 96-well Microfluor-1 U-bottom plate at a final concentration of 30 nM of each compound per well.

Hydrolysis reactions were initiated by the addition of 8.8 nM vinckepain-2 or 3.9 nM vinckepain-2$_{A23G}$ and monitored fluorimetrically with a Molecular Devices SpectraMax Gemini spectrophotometer, with excitation at 380 nm and emission at 460 nm. Assays were performed at 37 °C in 100 mM sodium acetate (pH 5.5)/100 mM NaCl/10 mM DTT/1 mM EDTA/0.01% Brij-35/1% (v/v) DMSO.

Assays of vinckepain-2 activity

Activity was routinely assessed by measurement of the hydrolysis of fluorogenic peptide-AMC substrate, as described previously [7]. For substrate gel analysis, samples were mixed with SDS/PAGE sample buffer lacking 2-mercaptoethanol and electrophoresed in a polyacrylamide gel, co-polymerized with 0.1% (w/v) gelatin [6]. The gel was then washed twice (30 min, 25 °C) with 2.5% (v/v) Triton X-100 and incubated overnight at 37 °C in 100 mM sodium acetate (pH 5.5)/10 mM DTT before staining with Coomassie Blue.

Substrate kinetics

The concentrations of vinckepain-2 and vinckepain-2$_{A23G}$ were determined by active-site titration with *trans*-epoxysuccinyl-L-leucilamido-(4-guanidino)butane (E-64). Rates of hydrolysis of peptide-AMC substrates in the presence of constant enzyme (2.5–25 nM) and different substrate concentrations were determined

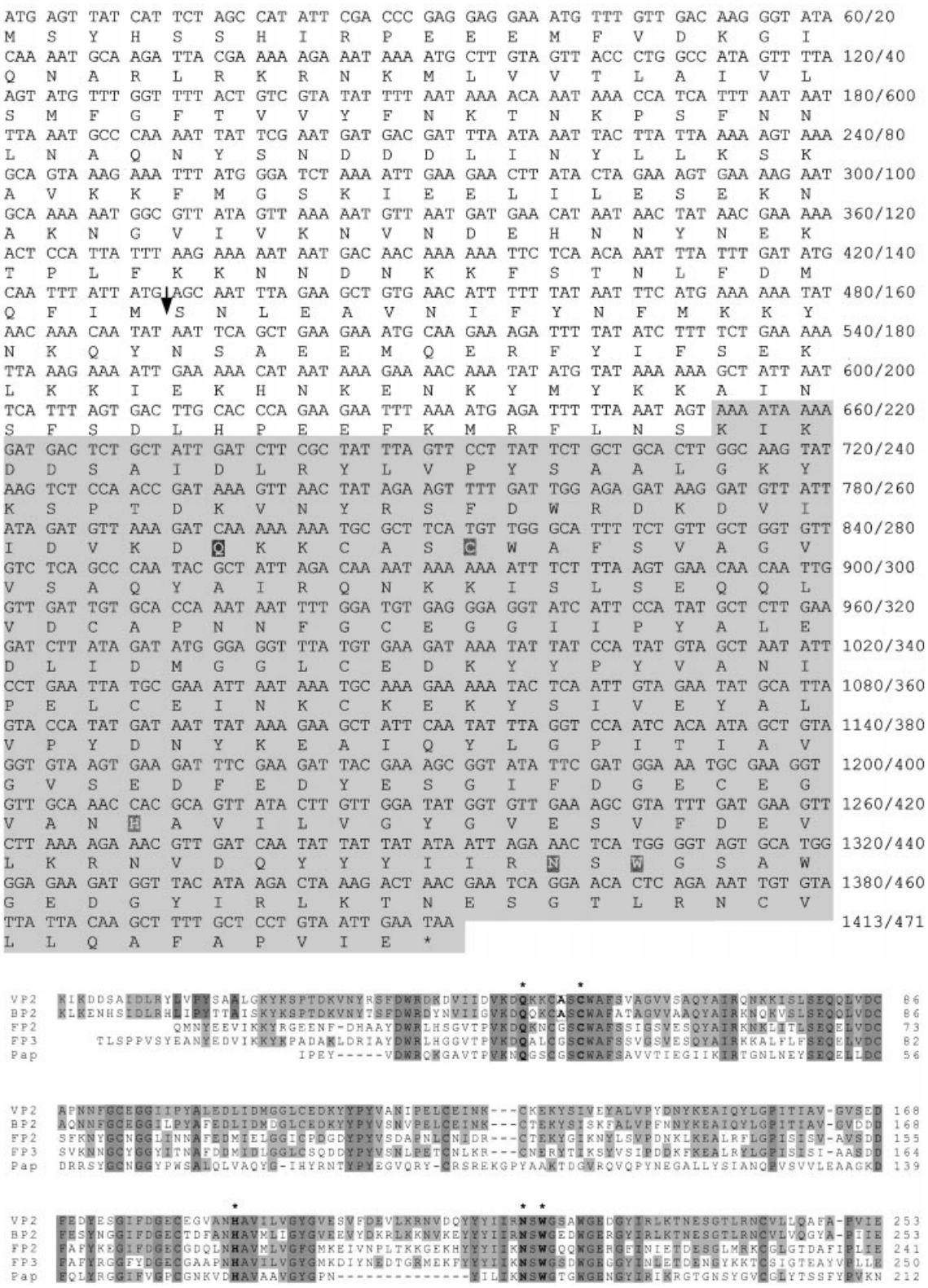

Figure 1 Nucleotide and deduced amino acid sequence of vinckepain-2 (VP2), and sequence alignments

Upper panel: the full coding sequence is shown. The ↓ indicates the N-terminus of the proenzyme expression construct. The mature protease domain is shaded, and conserved papain-family active-site residues are darkly shaded. Lower panel: the sequences of mature VP2, berghepain-2 (BP2), falcipain-2 (FP2), FP3 and papain (Pap) were aligned using the CLUSTAL W method. Numbering corresponds to the known or predicted (for BP2) mature cleavage site of each protease, dashes represent gaps required for optimal alignment, and * indicates conserved active-site amino acids. Amino acids that are identical in all five proteases are darkly shaded and identities with VP2 are lightly shaded. The unusual Ala[23] (papain numbering system) of VP2 and BP2 is boxed.

in 100 mM sodium acetate buffer (pH 5.5)/10 mM DTT. Fluorescence (excitation 355 nm; emission 460 nm) was continuously measured for 30 min (5 min for Z-Phe-Arg-AMC) at 25 °C. The kinetic constants K_m and V_{max} were determined using Prism (GraphPAD Software for Science, San Diego, CA, U.S.A.).

RESULTS

Identification of cysteine protease genes of rodent malaria parasites

A homologue of falcipain-1, now termed vinckepain-1, was identified previously [13]. Since it now appears that falcipain-2 and -3 are the principal haemoglobinases of *P. falciparum* [7,8,20], we were interested in identifying homologues of these enzymes. A BLAST search of available *Plasmodium* sequences (National Center for Biotechnology Information) identified a 602 bp *P. berghei*-expressed sequence tag, which was amplified and used to identify ORFs encoding very similar proteases of *P. berghei* and *P. vinckei*. Amplification and sequencing of each cDNA confirmed that these genes are transcribed in erythrocytic-stage parasites and that they do not contain introns. The ORFs encode fairly typical papain-family cysteine proteases (Figure 1). The predicted mature protease sequences are 78 % identical and most similar in sequence, among characterized proteins, to falcipain-2 and -3 (vinckepain-2 was 49 % identical with falcipain-2 and 46 % identical with falcipain-3). They also share a number of unusual features with the falcipains, as follows. (a) They contain unusually large pro-sequences [7,8]. (b) The pro-sequences predict membrane-spanning domains near the N-termini [7,8]. (c) As is the case for falcipain-2 and -3, the mature proteases are predicted to contain unusual, mature protease-processing sites, with cleavage adjacent to short N-terminal extensions, which are unique among reported papain-family proteases [7,8]. In falcipain-2, the 17-amino-acid extension mediated protein folding independent of the prodomain [21]. (d) Some unusual sequences near active-site residues were conserved among the plasmodial proteases [7,8]. The rodent plasmodial protease sequences differ from their *P. falciparum* homologues, however, in that they contain a substitution of Gly^{23} with Ala, which is highly unusual for papain-family enzymes [22]. This substitution is also present in homologous genes from two other rodent malaria parasites, *P. yoelii* (The Institute for Genomic Research) and *P. chabaudii* (Sanger Institute, Cambridge, U.K.).

Blots of restriction endonuclease-digested genomic *P. vinckei* and *P. berghei* DNA were probed with transcripts encoding the mature domains of the proteases. For each digested sample, a single fragment hybridized, indicating that vinckepain-2 and berghepain-2 are single-copy genes (Figure 2A) and, therefore, that in each species a single gene appears to be the closest orthologue to both falcipain-2 and -3. Faintly hybridizing bands probably represent more distant homologues (e.g. vinckepain-1 [13]). Northern blots of erythrocytic-stage *P. vinckei* and *P. berghei* RNA, probed with the protease genes, identified 2.3 kb transcripts (Figure 2B), indicating that the protease genes are transcribed in erythrocytic-stage parasites.

Expression, refolding and processing of recombinant proteases

A construct encoding portions of the prodomain and all of mature vinckepain-2 was amplified, cloned into the pQE-30 expression vector and used to transform *E. coli*. The protease was expressed as insoluble protein, solubilized and purified by Ni-NTA chromatography, and then refolding was optimized using a systematic microdilution approach as described previously for falcipain-2 [15]. The refolded protease was processed to an enzymically active species of the size predicted for the mature form on exposure to an acidic buffer (Figure 3A). The N-terminal sequence of refolded vinckepain-2 showed microheterogeneity, with identification of the sequences SKIKD- and KIKDD.

Immunoblots with an antiserum generated against vinckepain-2 and substrate gel analysis recognized a single active protease of identical size with recombinant vinckepain-2 in *P. vinckei*-soluble extracts (Figures 3B and 3C). Thus vinckepain-2 is translated and is enzymically active in erythrocytic-stage *P. vinckei* parasites. The recombinant enzyme had considerably greater activity than native vinckepain-2, although it is not clear if this difference was

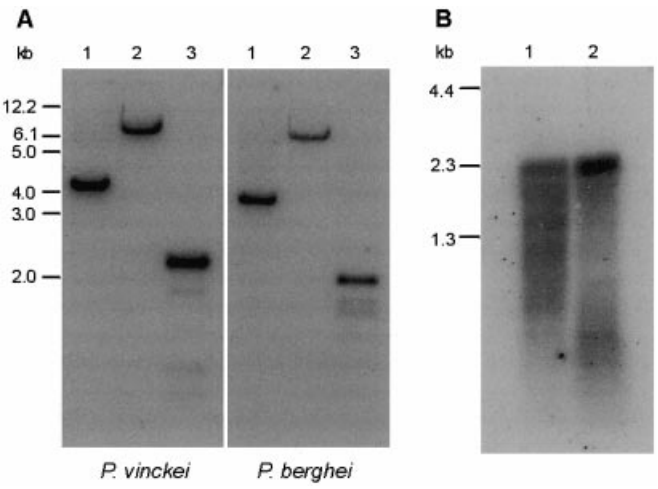

Figure 2 (A) Southern-blot and (B) Northern-blot analyses

(**A**) Genomic *P. vinckei* and *P. berghei* DNAs were digested with restriction endonucleases (lane 1, *Eco*RI; lane 2, *Eco*RV; lane 3, *Pac*I), electrophoresed on a 0.7 % agarose gel, transferred to nylon membranes and hybridized with specific probes. Membranes were washed under high-stringency conditions and evaluated by autoradiography. The positions of molecular-mass markers are shown in kb. (**B**) *P. vinckei* (lane 1) and *P. berghei* (lane 2) total RNAs were electrophoresed, transferred and hybridized with specific probes. Membranes were washed under low-stringency conditions and evaluated by autoradiography.

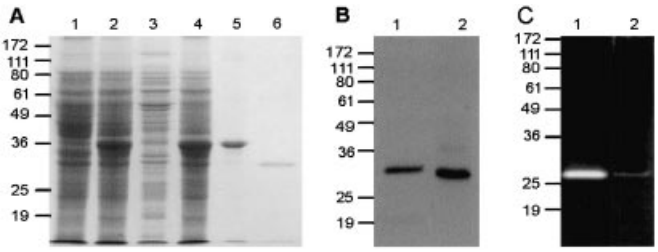

Figure 3 Expression and purification of recombinant vinckepain-2

(**A**) The −73VP2 fragment was expressed in *E. coli*, purified and refolded, and was evaluated by reducing SDS/PAGE (12% gel) and staining with Coomassie Blue. The lanes show non-induced cells (lane 1), cells induced with IPTG (lane 2), soluble (lane 3) and insoluble (lane 4) cellular fractions, material purified with Ni-NTA (lane 5), and protein remaining after refolding by the method described above (lane 6). (**B**, **C**) Recombinant vinckepain-2 (lane 1, equal quantities in the two gels) and a *P. vinckei* soluble extract (lane 2, equal quantities in the two gels) were resolved with reducing SDS/PAGE (**B**) and non-reducing gelatin substrate SDS/PAGE (**C**) and then blotted and reacted with anti-(vinckepain-2) antiserum (**B**) or stained for proteolytic activity (**C**). The positions of molecular-mass markers are shown in kDa.

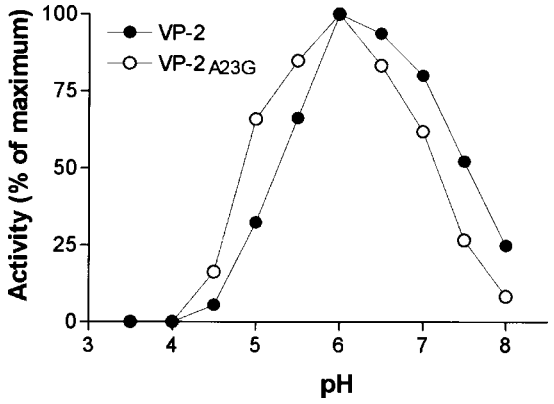

Figure 4 Optimum pH of recombinant proteases

The activities of recombinant vinckepain-2 (VP-2) (8.75 μM) and VP-2_{A23G} (9.8 μM) were assayed against 50 μM Z-Leu-Arg-AMC in 100 mM sodium acetate (pH 3.5–6.0) or sodium phosphate (pH 6.5–8.0), and 10 mM DTT. Release of AMC was continuously monitored for 30 min, slopes of fluorescence over time were calculated, and results were expressed as percentage of maximum activity.

Table 1 Substrate hydrolysis by plasmodial cysteine proteases

Values are means for an experiment performed in duplicate; individual results varied by less than 10%. Values for falcipain-2 and -3 were as reported previously [7,8], except for the substrate Z-Leu-Leu-Arg-AMC. n.d., not determined; n.c., not cleaved; Bz-, benzoyl; Boc-, *t*-butyloxycarbonyl; VP2, vinckepain-2.

	k_{cat}/K_m ($M^{-1} \cdot s^{-1}$)			
Substrate	VP2	$VP2_{A23G}$	Falcipain-2	Falcipain-3
Z-Leu-Arg-AMC	5.26×10^3	1.58×10^6	1.06×10^5	5.90×10^2
Z-Phe-Arg-AMC	9.70×10^1	*	4.51×10^4	1.20×10^2
Z-Val-Arg-AMC	1.40×10^2	6.63×10^4	5.44×10^3	2.60×10^1
Z-Leu-Leu-Arg-AMC	7.16×10^4	5.03×10^6	1.86×10^5	3.02×10^4
Z-Val-Leu-Arg-AMC	3.68×10^4	3.16×10^6	6.19×10^5	2.24×10^3
Z-Val-Val-Arg-AMC	3.23×10^4	3.69×10^6	4.03×10^3	1.99×10^2
Bz-Phe-Val-Arg-AMC	4.61×10^4	n.d.	n.c.	n.c.
Boc-Val-Leu-Lys-AMC	1.46×10^4	n.c.	4.97×10^4	6.23×10^2

* Could not be determined due to sigmoidal substrate curve.

due to biochemical differences between the enzymes, a difference in their stabilities or other factors.

Biochemical properties of recombinant vinckepain-2

Recombinant vinckepain-2 shared, with the falcipains and many other papain-family enzymes, a requirement of reducing conditions for maximal activity (maximal activity at 2 mM GSH or DTT) and an acidic pH optimum (Figure 4). Considering the hydrolysis of a selected set of synthetic peptide substrates, vinckepain-2 shared, with both falcipain-2 and -3, a preference for peptides with Leu at the P2 position, but kinetic constants varied markedly between the enzymes. In particular, vinckepain-2 had a marked preference for tripeptide (k_{cat}/K_m = 14000–72000 $\cdot M^{-1} \cdot s^{-1}$) over dipeptide ($k_{cat}/K_m < 100$–5000 $\cdot M^{-1} \cdot s^{-1}$) substrates and a small (< 2-fold) preference for Val over Phe at P2 (Table 1). Considering protease inhibitors, vinckepain-2 was fully inhibited by the standard cysteine protease inhibitors E-64 and leupeptin and not by inhibitors

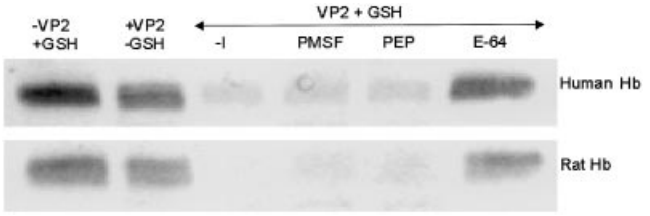

Figure 5 Haemoglobinase activity of vinckepain-2

Each reaction (30 μl) contained 3 μg of human or rat haemoglobin in 100 mM sodium acetate (pH 6.0) in the absence or presence of recombinant vinckepain-2 (VP-2) (600 nM), GSH (2 mM) and protease inhibitors (concentrations as in Figure 4), as labelled. The samples were incubated at 37 °C for 2 h and the reaction was stopped by the addition of SDS/PAGE loading buffer, and the remaining substrate was resolved by SDS/PAGE and Coomassie Blue staining. PEP, pepstatin.

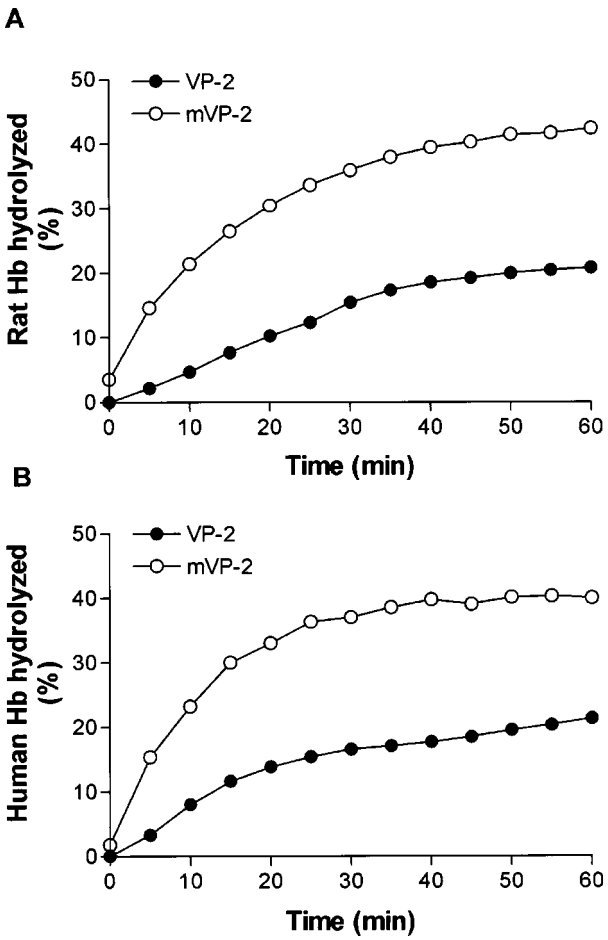

Figure 6 Haemoglobin hydrolysis by recombinant wild-type vinckepain-2 (VP-2) and mutant VP-2 (mVP-2)

Each reaction (500 μl) contained 50 μg of rat haemoglobin (**A**) or human haemoglobin (**B**) in 100 mM sodium acetate, 1 mM GSH (pH 5.5) at 37 °C in the absence (control) or presence of recombinant wild-type VP-2 or mVP-2 (200 nM). At each time point, absorbance *A* at 410 nm was determined, and the percentage of control absorbance was subtracted from 100 to yield the percentage of hydrolysis.

of other classes of proteases. However, the fluoromethyl ketone Z-Phe-Arg-CH_2F, a rapid and potent inhibitor of the falcipains and many other papain-family enzymes, was a relatively

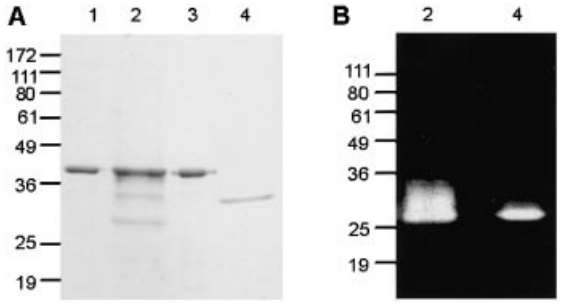

Figure 7 Refolding and processing of recombinant wild-type and mutant vinckepain-2 (VP2)

Equal quantities (2.5 mg) of Ni-NTA-purified recombinant VP2 and $VP2_{A23G}$ were incubated in 50 ml of the optimal refolding buffer at 4 °C for 21 h. Non-refolded Ni-NTA-purified VP2 (lane 1), $VP2_{A23G}$ (lane 3), refolded VP2 (lane 2) and VP_{A23G} (lane 4) were concentrated and resolved with reducing SDS/PAGE (12%) (**A**) and gelatin substrate non-reducing SDS/PAGE (**B**). The positions of molecular-mass markers are labelled in kDa.

poor inhibitor of vinckepain-2, consistent with the poor activity of this enzyme against the substrate Z-Phe-Arg-AMC (results not shown).

Vinckepain-2 is a haemoglobinase

Recombinant vinckepain-2 was incubated with human and rat haemoglobins. These substrates were hydrolysed in a manner that was dependent on the physiological reductant GSH and was inhibited by the cysteine protease inhibitor E-64, but not by inhibitors of Ser (PMSF) or Asp (pepstatin) proteases (Figure 5). Thus, as is the case with the falcipains, vinckepain-2 is an effective haemoglobinase.

Substitution of Ala^{23} markedly changes the activity of vinckepain-2

The Ala at position 23 in vinckepain-2 was substituted with Gly, the residue present at this location in the majority of reported papain-family proteases. The mutated enzyme (vinckepain-2_{A23G}) shared with the wild-type enzyme the requirement of reducing conditions for maximal activity, an acidic pH optimum and inhibition by standard cysteine protease inhibitors (Figure 4). Vinckepain-2_{A23G} also had preferences for tripeptide over dipeptide substrates and for Z-Leu-Arg-AMC over Z-Val-Arg-AMC and Z-Phe-Arg-AMC (Table 1). However, vinckepain-2_{A23G} was markedly more active than the wild-type enzyme against all tested substrates, with, in general, k_{cat}/K_m values about two orders of magnitude greater for the mutant enzyme. The activity of vinckepain-2_{A23G} against each tested substrate was also much greater than that of falcipain-2 or falcipain-3 (Table 1). Most of this improvement in activity was due to increased k_{cat}, although K_m values were also consistently lower for mutant vinckepain-2 (results not shown). Vinckepain-2_{A23G} was also more active against haemoglobin; in a spectrophotometric assay, both human and rat haemoglobins were hydrolysed much more rapidly by the mutant than by the wild-type enzyme (Figure 6).

Recombinant vinckepain-2_{A23G} also differed from the wild-type enzyme in that it was more effectively processed. As was the case with falcipain-2 [7], but not falcipain-3 [8], the two vinckepain-2 constructs were effectively folded and processed to active enzymes in alkaline refolding buffer (Figure 7). However, this process was more rapid with vinckepain-2_{A23G}, such that, after an overnight incubation in refolding buffer, vinckepain-2_{A23G}, but not the wild-type enzyme, was completely processed.

Use of peptide libraries to compare the extended substrate specificities of wild-type and mutant proteases

The activities of vinckepain-2 and vinckepain-2_{A23G} against two positional-scanning tetrapeptide substrate libraries were assessed (Figure 8). Results from the P4-P3-P2-diverse library (in which P1 was fixed for 20 different sublibraries) and results from the P4-P3-P2-P1-diverse library were very similar. The enzymes had a relative preference for positively charged amino acids at P1, as is common for papain-family proteases. Vinckepain-2 also readily hydrolysed substrates with a P1 Gly, an unusual activity for papain-family proteases that was known previously only for the plant protease glycyl endopeptidase [23]. Mutagenesis to vinckepain-2_{A23G} led to the more typical P1-specificity for Lys and Arg. Relative activity against substrates with Gly at P1 was much decreased. As is the case with falcipain-2 [7] and falcipain-3 [8], vinckepain-2 and vinckepain-2_{A23G} showed a preference for Leu at P2. These results are consistent with autohydrolysis of recombinant vinckepain-2 at the sequence LN↓SKIKD. Activity

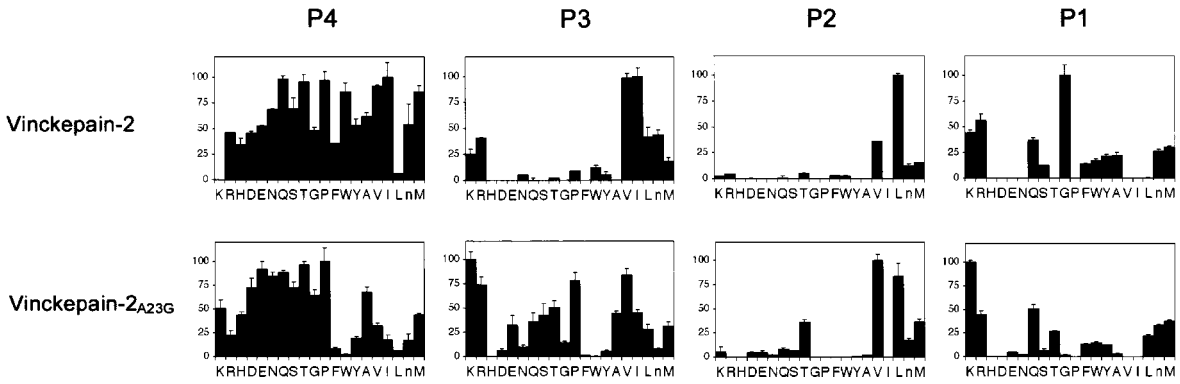

Figure 8 Specificity profiles of vinckepain-2 and vinckepain-2_{A23G} using positional-scanning synthetic substrate libraries

The P1-diverse library and P2, P3 and P4 libraries of the complete diverse library were used to determine specificities. Activities are displayed as percentage of the maximum for each position. Amino acids are represented by the single-letter code (N is norleucine). Error bars represent S.D. of results from three experiments.

against P2 Val substrates was also seen and it was greatest for vinckepain-2_{A23G}. Patterns of hydrolysis of substrates differing at P3 and P4 were complex. In summary, the comparison of hydrolysis patterns identified, most notably, the unusual ability of vinckepain-2 to cleave substrates with P1 Gly and the recovery of fairly typical papain-family P1-specificity after $Ala^{23} \rightarrow Gly$ mutagenesis.

DISCUSSION

We are interested in evaluating inhibitors of the *P. falciparum* haemoglobinases falcipain-2 and -3 as potential new anti-malarial agents. We have therefore identified homologous cysteine proteases from the rodent malaria parasites *P. vinckei* and *P. berghei*, and we expressed and characterized the *P. vinckei* enzyme. The *P. vinckei* cysteine protease will be a valuable reagent, since appreciation of its kinetic properties and inhibitor sensitivity will aid the interpretation of tests of the *in vivo* anti-malarial activity of cysteine protease inhibitors, which necessarily must be performed principally with non-primate malaria parasites. However, our characterization of this enzyme has shown that it differs significantly from related enzymes of *P. falciparum*, and differences between the cysteine proteases of rodent and human malaria parasites need to be considered to better interpret the results of *in vivo* screens.

Despite key differences, vinckepain-2 and berghepain-2 appear to be the closest orthologues of falcipain-2 and -3. Each of them appears to be a single counterpart for the two *P. falciparum* haemoglobinases, for the following reasons: they are both encoded by single-copy genes, PCR-based cloning approaches identified only one closely related protease gene in each species, and searches of nearly complete genomic databases of two other rodent malaria parasites, *P. yoelii* and *P. chabaudi*, also identified only one closely related gene in each species. All studied plasmodial species also encode a more distantly related papain-family enzyme (named falcipain-1 in *P. falciparum*), but this subfamily of enzymes is readily distinguished from the falcipain-2/3 subfamily based on overall amino acid identity and the presence of certain uniquely conserved sequences, e.g. $YYY^{169–171}$ (papain numbering system) is unique to the falcipain-2/3 subfamily [7,8,24]. Vinckepain-2 shared with the *P. falciparum* cysteine proteases an acidic pH optimum, preference for reducing conditions, optimal hydrolysis of peptide substrates with Leu at the P2 position and inhibition by typical cysteine protease inhibitors. It also demonstrated haemoglobinase activity at acidic pH, suggesting that it shares with falcipain-2 and -3 the functional importance as a food vacuole haemoglobinase [7,8].

Although the identification of falcipain-2/3 orthologues in rodent malaria parasites was not surprising, the expression and characterization of vinckepain-2 showed that it differed in key respects from the related enzymes of *P. falciparum*. Notably, vinckepain-2 differed in its substrate specificity, with remarkably low activity against the typical papain-family substrate Z-Phe-Arg-AMC. More detailed characterization of substrate specificity utilizing combinatorial peptide libraries demonstrated that, in contrast with most papain-family proteases, vinckepain-2 is capable of only very slow hydrolysis of substrates with Phe at the P2 position. Consistent with this result, vinckepain-2 was relatively weakly inhibited by Z-Phe-Arg-CH_2F. Our identification of unusual features of vinckepain-2 is consistent with previous studies, which showed unique substrate and inhibitor specificity for *P. vinckei* extracts [12] but were limited by the lack of purified enzymes for study. We now hypothesize that vinckepain-2 is the principal *P. vinckei* trophozoite cysteine protease and is a key haemoglobinase, and that differences between human and rodent parasite cysteine proteases explain, at least in part, relatively low *in vivo* anti-malarial efficacies of some cysteine protease inhibitors [11,12].

Although it has been identified as the *P. vinckei* enzyme most similar to falcipain-2 and -3, mature vinckepain-2 differs in sequence from the *P. falciparum* enzymes by approx. 50 %. Many non-conserved amino acids probably mediate differences in the biochemical activities of these proteases. However, the single substitution of Gly^{23} (papain numbering system) by Ala was noteworthy, since this amino acid appears to play a key role in catalysis and this is an uncommon substitution in papain-family proteases, where Gly^{23} is highly conserved. Indeed, substitutions at this position have been identified in proteases of four rodent malaria parasites but in only 14 of 157 papain-family (clan CA, subfamily C1A) sequences listed on the Merops protease database (http://www.merops.ac.uk/merops/ updated 13 February, 2002). Most of the other proteins with substitutions at position 23 have not been, to our knowledge, biochemically characterized. Only one listed sequence, from *Oryza sativa*, has Ala at this position, but the biochemical features of this putative protease are unknown. One protease of interest, with a substitution at position 23, is *Carica papaya* glycyl endopeptidase (papaya protease IV), an enzyme in which the substitutions $Gly^{23} \rightarrow Glu$ and $Gly^{65} \rightarrow Arg$ mediate unusual P1 substrate specificity, with a preference for substrates with P1 Gly, which is very unusual for papain-family enzymes [25–27].

Considering the importance of amino acid 23 in catalysis, we evaluated the effects of replacement of Ala^{23} with Gly, the amino acid at this position in nearly all characterized papain-family enzymes. This substitution led to three profound changes in vinckepain-2. Firstly, the substrate specificity of the mutant protease was altered. Vinckepain-2_{A23G} lost the strong preference for substrates with Gly at P1, which was a feature of the wild-type enzyme. Secondly, the mutant enzyme was remarkably active against peptide substrates. The activity of vinckepain-2_{A23G} against a number of substrates was consistently about two orders of magnitude greater than that of the wild-type enzyme and also consistently greater than those of falcipain-2 and -3. Finally, recombinant vinckepain-2_{A23G} more effectively underwent autohydrolysis than did the wild-type enzyme after overnight incubation in alkaline refolding buffer.

A detailed explanation of the changes engendered in vinckepain-2 by the $Ala_{23} \rightarrow Gly$ substitution will be unavailable until the structure of the enzyme is solved, but inferences can be made based on modelled structures and characterizations of related enzymes. The S1 subsite of papain-family proteases is bound by the catalytic Cys^{25} residue and, in most members of this family, Gly at positions 23 and 65. These amino acids create a large, accommodating pocket and exert relatively little influence on substrate specificity [25,27]. In glycyl endopeptidase, in contrast, Glu^{23} and Arg^{65} residues create a much more restrictive S1 pocket [25], leading to a strong preference for substrates with Gly at P1. Mutagenesis of cathepsin B, to replace the amino acids corresponding to Gly^{23} or Gly^{65} with those present in glycyl endopeptidase, also improved activity against substrates with Gly at P1, although a double mutant was enzymically inactive [26]. Vinckepain-2 and orthologues from other rodent malaria parasites show replacement of the amino acid corresponding to Gly^{23} with Ala, but the amino acid corresponding to Gly^{65} is conserved. Unlike the falcipains, which show a typical Arg/Lys P1 preference (Y. Choe, A. Singh, P. J. Rosenthal and C. S. Craik, unpublished work), vinckepain-2 preferred Gly at P1, but the single $Ala^{23} \rightarrow Gly$ mutation produced an enzyme with the typical P1 substrate preference for Arg and Lys, but not Gly. Modelling of vinckepain-2 did not clearly explain the profound

changes engendered by the $Ala^{23} \rightarrow Gly$ mutation (Y. Choe and C. S. Craik, unpublished work), but it suggested that the wild-type vinckepain-2 P1 pocket is more restricted, due to the presence of the Ala side chain, compared with most other papain-family proteases. However, as modelling did not explain fully the effects of mutagenesis, it is likely that substrate specificity is partly mediated by the influence of the substitution on the transition state of the enzyme.

It is less clear why the $Ala^{23} \rightarrow Gly$ mutation also dramatically altered the kinetics of hydrolysis of typical peptide substrates by vinckepain-2. We suggest that the mutation leads to a more flexible active-site backbone, allowing more rapid substrate turnover, and thus markedly improved k_{cat} against all tested substrates. The $Ala^{23} \rightarrow Gly$ mutation also improved the rate of folding and autoprocessing of vinckepain-2 to an active enzyme, probably due to an increased rate of autohydrolysis.

Why do falcipain-2/3 orthologues of rodent malaria parasites contain an unusual active-site substitution, which has a dramatic impact on both substrate specificity and enzyme activity? We hypothesize that a hypothetical 'parent' enzyme, equivalent to vinckepain-2_{A23G}, was so active as to offer an evolutionary disadvantage. The $Gly^{23} \rightarrow Ala$ mutation changes substrate specificity, but it also remarkably decreases activity against multiple substrates, and therefore, this change might have evolved to allow rodent malaria parasites to effectively hydrolyse necessary substrates (e.g. haemoglobin), without allowing deleterious uncontrolled proteolysis.

Our observation that an $Ala^{23} \rightarrow Gly$ substitution in vinckepain-2 markedly alters enzyme activity supports an essential role for this residue in mediating P1 substrate specificity and catalytic activity. Our studies have also highlighted important differences between cysteine protease haemoglobinases of human and rodent malaria parasites. As falcipain-2 and -3 are under consideration as chemotherapeutic targets, it is necessary to take into account differences between them and corresponding proteases of animal-model parasites when interpreting evaluations of potential anti-malarial protease inhibitors. Alternatively, our findings suggest that a genetic approach to anti-malarial drug discovery may be appropriate. In this case, rodent plasmodia containing mutations in cysteine proteases, rendering them more similar to falcipain-2 and -3, might be generated, and if these parasites remain viable, they would offer opportunities for screens of the *in vivo* activity of potential agents against *P. falciparum*-like rodent parasites.

This work was supported by grants from the National Institutes of Health (grant nos. AI35800 and RR01081 to P.J.R.; grant no. AI35707 to C.S.C.; Bethesda, MD, U.S.A.) and the Medicines for Malaria Venture. We thank Belinda Lee and Chi Chang for expert technical assistance.

REFERENCES

1 Breman, J. G. (2001) The ears of the hippopotamus: manifestations, determinants, and estimates of the malaria burden. Am. J. Trop. Med. Hyg. **64**, 1–11
2 Rosenthal, P. J. (2001) Protease inhibitors. In Antimalarial Chemotherapy: Mechanisms of Action, Resistance and New Directions in Drug Discovery, pp. 325–344, Humana Press, Totawa, NJ
3 Francis, S. E., Gluzman, I. Y., Oksman, A., Knickerbocker, A., Mueller, R., Bryant, M. L., Sherman, D. R., Russell, D. G. and Goldberg, D. E. (1994) Molecular characterization and inhibition of a *Plasmodium falciparum* aspartic hemoglobinase. EMBO J. **13**, 306–317
4 Hill, J., Tyas, L., Phylip, L. H., Kay, J., Dunn, B. M. and Berry, C. (1994) High level expression and characterization of Plasmepsin II, an aspartic proteinase from *Plasmodium falciparum*. FEBS Lett. **352**, 155–158
5 Banerjee, R., Liu, J., Beatty, W., Pelosof, L., Klemba, M. and Goldberg, D. E. (2002) Four plasmepsins are active in the *Plasmodium falciparum* food vacuole, including a protease with an active-site histidine. Proc. Natl. Acad. Sci. U.S.A. **99**, 990–995
6 Rosenthal, P. J., McKerrow, J. H., Aikawa, M., Nagasawa, H. and Leech, J. H. (1988) A malarial cysteine proteinase is necessary for hemoglobin degradation by *Plasmodium falciparum*. J. Clin. Invest. **82**, 1560–1566
7 Shenai, B. R., Sijwali, P. S., Singh, A. and Rosenthal, P. J. (2000) Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of *Plasmodium falciparum*. J. Biol. Chem. **275**, 29000–29010
8 Sijwali, P. S., Shenai, B. R., Gut, J., Singh, A. and Rosenthal, P. J. (2001) Expression and characterization of the *Plasmodium falciparum* hemoglobinase falcipain-3. Biochem. J. **360**, 481–489
9 Eggleson, K. K., Duffin, K. L. and Goldberg, D. E. (1999) Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite *Plasmodium falciparum*. J. Biol. Chem. **274**, 32411–32417
10 Rosenthal, P. J., Wollish, W. S., Palmer, J. T. and Rasnick, D. (1991) Antimalarial effects of peptide inhibitors of a *Plasmodium falciparum* cysteine proteinase. J. Clin. Invest. **88**, 1467–1472
11 Rosenthal, P. J., Lee, G. K. and Smith, R. E. (1993) Inhibition of a *Plasmodium vinckei* cysteine proteinase cures murine malaria. J. Clin. Invest. **91**, 1052–1056
12 Olson, J. E., Lee, G. K., Semenov, A. and Rosenthal, P. J. (1999) Antimalarial effects in mice of orally administered peptidyl cysteine protease inhibitors. Bioorg. Med. Chem. **7**, 633–638
13 Rosenthal, P. J. (1993) A *Plasmodium vinckei* cysteine proteinase shares unique features with its *Plasmodium falciparum* analogue. Biochim. Biophys. Acta **1173**, 91–93
14 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**, 4673–4680
15 Sijwali, P. S., Brinen, L. S. and Rosenthal, P. J. (2001) Systematic optimization of expression and refolding of the *Plasmodium falciparum* cysteine protease falcipain-2. Protein Expr. Purif. **22**, 128–134
16 Sasaki, K., Taura, F., Shoyama, Y. and Morimoto, S. (2000) Molecular characterization of a novel β-glucuronidase from *Scutellaria baicalensis georgi*. J. Biol. Chem. **275**, 27466–27472
17 Hu, Y. K., Eisses, J. F. and Kaplan, J. H. (2000) Expression of an active Na, K-ATPase with an α-subunit lacking all twenty-three native cysteine residues. J. Biol. Chem. **275**, 30734–30739
18 Shenai, B. R. and Rosenthal, P. J. (2002) Reducing requirements for hemoglobin hydrolysis by *Plasmodium falciparum* cysteine proteases. Mol. Biochem. Parasitol. **122**, 99–104
19 Harris, J. L., Backes, B. J., Leonetti, F., Mahrus, S., Ellman, J. A. and Craik, C. S. (2000) Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. Proc. Natl. Acad. Sci. U.S.A. **97**, 7754–7759
20 Rosenthal, P. J., Sijwali, P. S., Singh, A. and Shenai, B. R. (2002) Cysteine protease of malaria parasites: targets for chemotherapy. Curr. Pharm. Des. **8**, 1659–1672
21 Sijwali, P. S., Shenai, B. R. and Rosenthal, P. J. (2002) Folding of the *Plasmodium falciparum* cysteine protease falcipain-2 is mediated by a chaperone-like peptide and not the prodomain. J. Biol. Chem. **277**, 14910–14915
22 Berti, P. and Storer, A. C. (1995) Alignment/phylogeny of the papain superfamily of cysteine proteases. J. Mol. Biol. **246**, 273–283
23 Buttle, D. J. (1998) Glycyl endopeptidase. In Handbook of Proteolytic Enzymes (Barrett, A. J., Rawlings, N. D. and Woessner, J. F., eds.), pp. 562–565, Academic Press, San Diego, CA
24 Rosenthal, P. J. (1996) Conservation of key amino acids among the cysteine proteinases of multiple malarial species. Mol. Biochem. Parasitol. **75**, 255–260
25 O'Hara, B. P., Hemmings, A. M., Buttle, D. J. and Pearl, L. H. (1995) Crystal structure of glycyl endopeptidase from *Carica papaya*: a cysteine endopeptidase of unusual substrate specificity. Biochemistry **34**, 13190–13195
26 Fox, T., Mason, P., Storer, A. C. and Mort, J. S. (1995) Modification of S1 subsite specificity in the cysteine protease cathepsin B. Protein Eng. **8**, 53–57
27 Buttle, D. J. (1994) Glycyl endopeptidase. Methods Enzymol. **244**, 539–555

Received 10 May 2002/24 July 2002; accepted 8 August 2002
Published as BJ Immediate Publication 8 August 2002, DOI 10.1042/BJ20020753