Substrate mapping and inhibitor profiling of falcipain-2, falcipain-3 and berghepain-2: implications for peptidase anti-malarial drug discovery

Manoj K Ramjee, Nicholas S Flinn, Tracy P Pemberton, Martin Quibell, Yikang Wang, John P Watts

Biochemical Journal Immediate Publication. Published on 15 June 2006 as manuscript BJ20060422

Substrate mapping and inhibitor profiling of falcipain-2, falcipain-3 and berghepain-2: implications for peptidase anti-malarial drug discovery

Manoj K. Ramjee*, Nicholas S. Flinn, Tracy P. Pemberton, Martin Quibell, Yikang Wang, John P. Watts

Amura Therapeutics Limited, Horizon Park, Barton Road, Comberton, CB3 7AJ, U.K.

*To whom correspondence should be addressed (manoj.ramjee@amura.co.uk)

Tel: +44(0)1223-265265
FAX: +44(0)1223-265266

Short title: Malaria cysteine peptidase substrate and inhibitor mapping

Key words: malaria, cysteine protease, substrate, inhibitor, mapping

SYNOPSIS

The *Plasmodium falciparum* cysteine peptidases falcipain-2 (FP-2) and falcipain-3 (FP-3), which are members of the papain-like CAC1 family, are essential haemoglobinases and therefore potential anti-malarial drug targets. To facilitate a rational drug discovery programme, we analysed the synthetic substrate and model inhibitor profiles of FP-2 and FP-3 as well as berghepain-2 (BP-2); an ortholog from the rodent parasite *P. berghei*. With respect to substrate catalysis, FP-2 exhibited a promiscuous substrate profile based around a consensus non-primeside motif, FP-3 was somewhat more restricted and BP-2 was comparatively specific. Substrate turnover for FP-2 was driven by a basic or acidic P1 residue, whereas for FP-3 turnover occurred predominately through a basic P1 residue only and for BP-2 turnover was again mainly through a basic P1 residue for some motifs and surprisingly glycine in the P1 position for other motifs. Within these P1 binding elements, additional recognition motifs were observed with subtle nuances that switched substrate turnover on or off through specific synergistic combinations. The peptidases were also profiled against reversible and irreversible cysteine peptidase inhibitors, the results re-iterated the contrasting kinetic behaviour of each peptidase as observed through the substrate screens. The results showed that the substrate and inhibitor preferences of BP-2 were markedly different from those of FP-2 and FP-3, which when compared to each other also displayed similarities and some significant differences. In conclusion, the *in vitro* data highlights the current difficulties faced by a peptidase directed anti-malarial medicinal chemistry program where compounds need to be identified with potent activity against at least three peptidases, each of which displays distinct biochemical traits.

Abbreviations used: Abz, 2-amino benzoic acid; AMC, 7-amino-4-methyl coumarin; AFC, 7-amino-4-trifluoromethyl coumarin; DMF, dimethylformamide; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; FRET, fluorescence resonance energy transfer; HBTU, 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HOBt, 1-hydroxybenzotriazole; MES, 2-(N-morpholino)ethansulfonic acid; Mu-Phe-hPhe-FMK, *N*-methoxysuccinyl-phenylalanyl-homophenylalanyl-fluoromethyl ketone; Mu-Val-hPhe-FMK, *N*-methoxysuccinyl-valyl-homophenylalanyl-fluoromethyl ketone; Nal, 2-napthylalanine; Nle, norleucine; NMM, *N*-methylmorpholine; Thi, 2-thienylalanine; RFU, relative fluorescence units

INTRODUCTION

Malaria, a disease caused by several species of obligate protozoan parasites, remains one of the most prevalent and persistent diseases to affect the global population; with hundreds of millions of people infected and approximately 1-2 million deaths each year [1]. This situation is compounded by the increasing resistance of the malaria parasites, particularly *Plasmodium falciparum*, to conventional drugs. There is therefore great need to identify new therapeutically amenable targets and/or mechanisms with a view to developing small molecule inhibitors and/or vaccines to treat malaria [1].

Malaria parasites utilise a wealth of processes to complete the various stages of their lifecycle [1], for which a range of peptidases are essential for the completion of one or more of these stages [2]. In the human host, processes such as erythrocyte invasion, erythrocyte rupture upon escape as well as nutrient acquisition by degradation of host proteins, particularly haemoglobin, rely on the action of peptidases [3]. Early studies demonstrated the importance of these peptidases, since in the presence of peptidase inhibitors parasite development could be arrested whilst removal of inhibitor resulted in resumption of development. Biochemical studies, complemented by genomic [4] and proteomic studies, have enabled temporal [2] and spatial identification of the various peptidases to be made, allowing their precise biological role(s) to be addressed. To date members of the aspartic, metallo, serine and cysteine peptidase classes have been described from at least one or more species; a subject which has been reviewed elsewhere [3].

The papain-like cysteine peptidases derived from *P. falciparum,* the most prolific human malaria parasite, are termed falcipains (FP) and comprise four peptidases- FP: 1 [5], FP-2 [6], FP-2' [7] and FP-3 [8]. FP-1, located on chromosome 14, shares ~38-40% sequence identity with FP-2 and FP-3, and has been implicated in erythrocyte invasion by merozoites [9]. Gene disruption studies suggest that FP-1 is not essential to the erythrocytic stage of *P. falciparum* [5, 10], although gene silencing studies have alluded to a functional biological role for FP-1 [9, 11]. FP-2 and FP-2', which are ~96% identical, share ~68% identity with FP-3 and are encoded on chromosome 11. FP-2, FP-2' and FP-3 are food vacuole haemoglobinases [6, 8, 12] and their joint expression, in concert with aspartic peptidases, seems key to efficient hydrolysis [8]. All three enzymes have been cloned, over-expressed and biochemically characterised [6, 8, 12-14]. Gene disruption [15], coupled with inhibitor studies [3], have

established a critical role for FP-2 in haemoglobin hydrolysis by *P. falciparum* and as such it has been pursued as an anti-malarial target. However it seems that functional redundancy exists between the various falcipain peptidases and therefore, attention has to be paid to more than one enzyme during drug discovery and development [12, 16]. Coupled to this, the advancement of compounds inevitably requires demonstration of efficacy, firstly in cell-based assays and then in a relevant disease-related animal model [17]. It is here that cysteine peptidase anti-malarial drug discovery faces a conundrum. The initial animal models of choice are rodent-based; however non-engineered rodents do not support infection by human *P. falciparum* [18]. Instead, rodents are infected by various species-specific malaria parasites which are employed to model the human disease [17]. The *in vitro* and *in vivo* screening cascades are further complicated since the analogous rodent parasite infected cell-based assays are not currently viable. Thus a rational drug discovery design programme which utilises optimisation of compounds against falcipains both *in vitro* then through cell-based assays of infection, is subsequently required to show efficacy in rodent-based models where the target peptidase is functionally analogous, but may be biochemically distinct [14]. This potential dilemma has been highlighted previously, prompting the engineering of an immuno-compromised mouse capable of hosting *P. falciparum* infection [18]. Although a major step forward in peptidase anti-malarial discovery, this model is yet to receive widespread application. Reported herein is an extensive examination of the synthetic substrate and literature CAC1 peptidase inhibitor profiles for FP-2, FP-3 and berghepain-2 (BP-2). The results clearly indicated that, although FP-2 and FP-3 shared some common features, there were both subtle and significant differences in their biochemical behaviour. Moreover, the data also highlighted that BP-2 was markedly different from FP-2 and FP-3. Thus, the difficulties faced by a medicinal chemistry and drug discovery program are clearly evident even at the initial *in vitro* screening stages. The results of the current study are important if the intent is to develop approaches to simultaneously target both FP-2 and FP-3, and yet demonstrate efficacy by targeting BP-2, a related but distinct enzyme. The data suggests that it would be inappropriate to use the rodent *in vivo P. berghei* model for testing falcipain inhibitors unless transgenic parasites, in which the BP-2 gene had been replaced by the falcipain genes, can be used.

EXPERIMENTAL

Materials

Unless otherwise stated, all general chemicals and biochemicals were purchased from Sigma Chemical Company, Poole, Dorset, U.K., Fisher Scientific U.K., Loughborough, Leicestershire, U.K. Detergents (*e.g.* CHAPS, zwittergents, *etc.*) were purchased from Merck Biosciences UK, Beeston, Nottinghamshire, U.K. Stock solutions of substrate or inhibitor were made up to 10 mM in 100 % dimethylsulfoxide (DMSO) (Rathburns, Glasgow, U.K.) and diluted in buffer as appropriately required. In all cases the DMSO concentration in the assays was maintained at less than 1% (vol./vol.).

Solid phase peptide synthesis

All solid phase synthesis was performed using an "Fmoc/tBu" procedure [19] using standard solid phase synthesis resin washing protocols. All solvents were purchased from ROMIL Ltd. (Waterbeach, Cambridge, U.K.) at SpS or 'Hi-Dry' grade unless otherwise stated. General peptide synthesis reagents, amino acids and/or derivatives were purchased from Chem-Impex International Inc. (Wood Dale, IL 60191, U.S.A.), Bachem U.K. (St. Helens, Merseyside, U.K.), Merck Biosciences or Neosystems (Strasbourg, France) and were of the L-configuration unless otherwise stated. Peptide synthesis was carried out in repetitive cycles consisting of a coupling step, a reagent wash step, a piperidine:DMF (20:80) Fmoc de-protection step, then a further extensive wash step followed by the next coupling round. Between each step excess reagent and solvent were removed by application of a vacuum. Each coupling step was initiated by activating a five-mole excess of Fmoc-amino acid (with respect to total solid phase loading capacity) in the presence of a five-mole excess HBTU (Merck Biosciences), a five-mole excess HOBt (Merck Biosciences) and a ten-mole excess of NMM (Aldrich).

Synthesis of the spatially addressed FRET substrate library synthesis

The FRET substrate library was based on the utilisation of the ABZ/Tyr(3-NO_2) FRET pair [20]. Solid phase library syntheses were performed using gears (Mimotope SynPhase GAP 1.3 µmol; Perbio Sciences U.K. Ltd., Tattenhall, Cheshire, U.K.) pre-derivatised with RINK-amide linker, in an eight by twelve 96-well microtitre plate format following standard coupling, washing, deprotection and

cleavage protocols [21]. Peptide synthesis was based on the manufacturer's protocols (Protocol STN 002-2; http://www.synphase.com/). Using the cyclic repetitive protocol described above, gears were elaborated with coupling cycles of Fmoc-Glu-OH followed by Fmoc-Tyr(3-NO_2)-OH (ChemImpex) to yield NH_2-Tyr(3-NO_2)-Glu-solid phase. Library construction was carried out using a split-and-mix strategy comprising of four rounds of synthesis (Supplementary Figure 1).

Large-scale peptide synthesis

Large-scale peptide syntheses were performed manually using lanterns (Mimotope SynPhase GAP 8 µmol; Perbio Sciences U.K. Ltd., Tattenhall, Cheshire, U.K.) pre-derivatised with RINK-amide linker, following standard coupling, washing, deprotection and cleavage protocols [21]. Upon completion of the sequence, crude product was cleaved with a cocktail of 92.5% TFA:2.5% triisopropylsilane:2.5% water:2.5% ethanedithiol for 90 min. after which the resin was removed by filtration and the filtrate concentrated by sparging with nitrogen gas. The crude products were precipitated by addition of 50 ml cold methyl tert-butyl ether (MTBE) and precipitates collected by centrifugation (4960 *g* for 5 min). The supernatant was discarded and the process repeated. The final precipitate was solubilised using HPLC buffer and purified as required.

Inhibitor compound synthesis

N-(1-(cyanomethylcarbamoyl)cyclohexyl)-4-(1-(2-methoxyethyl)piperidin-4-yl) benzamide (**1**) [22], (*S*)-*N*-(1-(cyanomethylamino)-4-methyl-1-oxopentan-2-yl)biphenyl-4-carboxamide (**2**) [23], *N*-((*S*)-4-methyl-1-oxo-1-((*R*)-3-oxo-1-(pyridin-2-ylsulfonyl)azepan-4-ylamino)pentan-2-yl)benzofuran-2-carboxamide (**3**) [23], *N*-((*S*)-4-methyl-1-oxo-1-((*S*)-3-oxo-1-(pyridin-2-ylsulfonyl)azepan-4-ylamino)pentan-2-yl) benzofuran-2-carboxamide (**4**) [23], benzyl (2*S*,2'*S*)-1,1'-(2,2'-carbonylbis(hydrazine-2,1-diyl))bis(4-methyl-1-oxopentane-2,1-diyl)dicarbamate (**5**) [24], Cbz-Leu-NH-NH-C(O)-O-Bz (**6**) [25], Cbz-Leu-NH-NH-C(O)-O-Ph (**7**) [25] and N-((S)-4-methyl-1-oxo-1-((S,E)-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-ylamino)pentan-2-yl) morpholine-4-carboxamide (LHVS) (**12**) [26] were synthesised using previously published methods.

HPLC and HPLC-MS analysis

In all cases, solvent A consisted of 0.1% (v/v) trifluoroacetic acid (TFA) in water and solvent B consisted of 90% acetonitrile mixed with 10% solvent A. Analytical HPLC was conducted using a Jupiter C4 column (5µ, 300Å, 250 x 4.6 mm; Phenomenex, Macclesfield, Cheshire, U.K.) on a manual HP1100 system (Agilent Technologies, Bracknell, U.K.) with data collection at 215 nm. Unless otherwise stated, a linear increasing gradient of 10 – 90% B over 25 min. at 1.5 ml/min. was used for column elution. Semi-preparative HPLC purification of crude samples was performed on a Jupiter C4 column (250 x 10 mm; 5µ, 300 Å; Phenomenex) using a linear increasing gradient of solvent B in solvent A using empirically derived gradients at 4.0 ml/min. The eluant absorbance was monitored at 230 nm and desired fractions analysed and then lyophilised (Heto Holten; Thermo Electron Corp., Basingstoke, U.K.).

HPLC-MS analysis was performed on Agilent HP1100 series LC/MSD with a linear increasing gradient of 10 – 90% B in A over 10 min. on a Columbus C8 column (5µ, 300Å, 50 x 2.0 mm; Phenomenex) at 0.6 ml/min. Absorbance data were collected using diode array detector and spectra for all peaks were stored as part of the chromatogram. The mass spectrometer was set to API-ES ionisation mode, positive polarity; scanning in the 200-2000 Da mass range with a gas temperature set to 350°C, nebuliser pressure of 60 p.s.i and drying gas flow set at 12 L/minute.

The cleavage sites for the FRET peptides were identified by incubating substrate (100 µM) with enzyme at room temperature for 3-6 h prior to analysis by HPLC-MS.

Peptidase activity assays

Unless otherwise stated, all kinetic experiments were performed at 25±0.5°C. Fluorescence high throughput assays were carried out in either 384-well microtitre plates (Corning Costar 3705 plates, Fisher Scientific) or 96-well 'U' bottomed Microfluor W1 microtitre plates (Thermo Labsystems, Ashford, Middlesex, U.K.). Fluorescence assays were monitored using a SpectraMax Gemini fluorescence plate reader (Molecular Devices, Crawley, Sussex, U.K.). Absorbance assays were carried out in flat-bottomed 96-well plates (Spectra; Greiner Bio-One Ltd., Stonehouse, Gloucestershire, U.K.) using a SpectraMax PLUS384 plate reader (Molecular Devices). In all cases, SOFTmax Pro software version 3.1.2 was used for data collection and rate analysis. For substrates employing the coumarin fluorophores, assays were monitored at an excitation wavelength of 365 nm and an emission wavelength of 450 nm and the fluorescence plate reader calibrated with AMC in assay

buffer. For substrates employing a 2-amino-benzoyl (Abz) fluorophore, assays were monitored at an excitation wavelength of 310 nm and an emission wavelength of 445 nm; the fluorescence plate reader was calibrated with 2-amino-benzamide (Fluka) in assay buffer. Unless otherwise indicated, all the peptide-coumarin substrates were purchased from Bachem UK. Z-LR-AMC and Z-VLR-AMC were synthesised by Amura Therapeutics Limited (see Supplementary Information). The FRET peptide concentrations were calibrated at 445 nm in 20 mM sodium phosphate, pH 9.0 using the nitro-tyrosine chromophore [20].

Recombinant FP-2, FP-3 and BP-2 [6, 8, 27] were obtained from Prof. P. J. Rosenthal, University of California, San Francisco, U.S.A. In brief, proteins were produced in *E. coli*, solubilised from inclusion bodies, purified by nickel-nitrilotriacetic acid chromatography, refolded, and further purified by anion exchange (Q-Sepharose) chromatography. Purity was demonstrated by the detection of a single protein band after SDS-PAGE and staining with Coomassie Blue [28]. The active site titrations of FP-2 and FP-3 were carried out using Mu-Phe-hPhe-FMK (Sigma) whereas the active titration of BP-2 was carried out using E-64 (Bachem). Routine peptidase assays were carried out in 100 mM sodium acetate; pH 5.75 containing 1 mM EDTA and 10 mM L-cysteine. Peptidase activity assays were routinely monitored for 30 min. and the rates determined by the linear increase in fluorescence signal over time. Initial velocities were used for calculations of the kinetic constants and steady-state velocities for calculations of the inhibition constants. Data were routinely analysed using Prism ver 4.03 (GraphPad Inc, San Diego, CA, U.S.A.).

Cathepsin K, B, S (Merck Biosciences), V (R&D Systems, Oxford, U.K.,) and L (Athens Research and Technology, Athens, GA, U.S.A.) were assayed according to literature methods [29].

FRET-based substrate library screening

For high-throughput screening of the FRET substrate library, assays were routinely carried out by dispensing 10 µl of assay buffer followed by 1 µl of a 250 µM solution of the substrate library diluted in 1 mM sodium acetate; pH 5.75. Assays were initiated by the addition of 10 µl enzyme in assay buffer. Peptidase activity was monitored for 30 min. and the rate determined by the increase in fluorescence signal over time. SigmaPlot ver 8.0 (Systat Software UK Limited, London, U.K.) was used to plot the rate data as a colour density array utilising the colour transition toolbox.

Each peptidase dataset was treated independently of the other with zero rates set as black and the maximum observed rate as yellow.

Measurement of the inhibition constants

The inhibition constant (K_i) for test compounds was determined on the assumption that inhibition was reversible and occurred by a pure-competitive mechanism (Supplementary Figure 2). The K_i values were calculated from the dependence of enzyme activity as a function of inhibitor concentration, by direct non-linear regression analysis (Prism v 4.03) using the equation $v_i=(V_{max}.[S_o])/([S_o]+(K_m.(1+[I]/K_i)))$ [30]; where '$\boldsymbol{v_i}$' was the observed residual activity, '$\boldsymbol{V_{max}}$' was the observed maximum activity (*i.e.* in the absence of inhibitor), '$\boldsymbol{K_m}$' was the apparent macroscopic binding constant for the substrate, '**[S_o]**' was the initial substrate concentration, '$\boldsymbol{K_i}$' was the dissociation constant and '**[I]**' was the inhibitor concentration.

In situations where the dissociation constant (K_i) approached the enzyme concentration, the K_i values were calculated using a quadratic solution in the form described by $v_i=(F(E_o-I_o-K_i^{app}+\sqrt{((E_o-I_o-K_i^{app})^2+4.K_i^{app}.E_o)}))/2$ [31]. In this case '$\boldsymbol{v_i}$' was the observed residual activity, '**F**' was the difference between the maximum activity (*i.e.* in the absence of inhibitor) and minimum enzyme activity, '$\mathbf{E_o}$' was the total active enzyme concentration, '$\boldsymbol{K_i^{app}}$' was the apparent dissociation constant and '$\mathbf{I_o}$' as the inhibitor concentration. Curves were fitted by non-linear regression analysis (Prism) using a fixed value for the enzyme concentration. The relationship $K_i^{app}= K_i.(1+[S_o]/K_m)$ was used to account for the substrate kinetics, where '$\boldsymbol{K_i}$' was the inhibition constant.

Determination of the rate of reaction for irreversible inhibitors

Where applicable, the concentration dependence of the observed rate of reaction (k_{obs}) of ligand with enzyme was analysed by determining the rate of enzyme inactivation under pseudo-first order conditions in the presence of substrate. Assays were carried out by the addition of various concentrations of inhibitor to assay buffer containing substrate. Assays were initiated by the addition of enzyme to the reaction mixture and the change in fluorescence monitored over time. During the course of the assay (60 min.) less than 10% of the substrate was consumed. The activity fluorescence

progress curves were fitted by non-linear regression analysis (Prism) using the equation $F=v_s t + ((v_o-v_s).(1-e^{(k_{obs}.t)})/k_{obs} + D$ [32]; where '**F**' was the fluorescence response, '**t**' was time, '$\boldsymbol{v_o}$' was the initial velocity, '$\boldsymbol{v_s}$' was the equilibrium steady-state velocity, '$\boldsymbol{k_{obs}}$' was the observed pseudo first-order rate constant and '**D'** was the intercept at time zero (*i.e.* the ordinate displacement of the curve). The second order rate constant was obtained from the slope of the line of a plot of k_{obs} versus the inhibitor concentration (*i.e.* $k_{obs}/[I]$). The equation $k_{inact}= (k_{obs}(1+[S_o]/K_m))/[I]$ was used to correct for substrate presence [31, 32].

RESULTS

Optimisation of the peptidase activity assay condition

In order to facilitate the screening effort, robust assays for each peptidase, amenable to automation and high throughput format, were required. Using literature-based methods [6, 8, 12, 14], we screened a set of approximately eleven hundred assay conditions encompassing different buffers (*e.g.* phosphate, bis-tris propane, *etc*.), various buffer compositions (*e.g.* mixed buffers like acetic acid/MES/tris), pH ranging from pH 4.5 to 10 as well as the presence of a range of additives (*e.g.* detergents, polyethylene glycol, glycerol, and alkali salts). In all three cases, the highest peptidase activity was observed in the range pH 5-6, with activity decreasing sharply above pH 7.5. Although a number of diverse buffer systems could be accommodated as determined by enzyme activity, no significantly better assay buffer system than that previously reported [6, 8] was found. In contrast to some cathepsins, the presence of detergent(s) did not significantly enhance peptidase activity. Although the inclusion of DTT in the assay buffer produced higher protease activity at relatively lower concentration; the use of higher concentrations of L-cysteine or 2-mercaptoethanol could compensate for this. We decided to employ L-cysteine as a reductant since, under prolonged incubation, the dithiane oxidation product of DTT [33] quenched the absorbance of the Abz chromophore; especially considering that DTT was employed at millimolar concentrations and the FRET substrates were at micromolar concentrations. Based on these results, the standard assay buffer employed for monitoring all three peptidase activities was 100 mM sodium acetate, pH 5.75 containing 1 mM EDTA and 10 mM L-cysteine.

Substrate screen

AMC-based substrate screen

In order to quickly and efficiently establish an activity assay, a set of fifty seven peptide-AMC substrates was screened against all three proteases. Coumarin-based substrates provide a sensitive method by which peptidase activity can be monitored. The synthesis of coumarin-based substrates in combinatorial library format has been reported [34] and positional scanning AMC library screens have been reported previously for vivapain-2 and vivapain-3 [35] as well as vinckepain-2 [27] along with a wealth of data for individual AMC substrates against FP-2 and FP-3, and to a lesser

extent BP-2 [6, 8, 12, 14]. The AMC screening data (Figure 1) are in general agreement with those previously reported but also highlight a number of novel peptide-AMC substrates. The data demonstrated common and unique substrate preferences between the three enzymes. On the basis of the peptide-AMC screening data, with substrates screened at a fixed concentration (10 μM) against all three peptidases, the specific activity of FP-2 was approximately one hundred times greater than that of FP-3 and approximately ten-fold greater than that of BP-2. Substrates utilised by all peptidases included Z-LR-AMC, Z-VLR-AMC and boc-VLK-AMC. Z-LR-AMC was most preferred by FP-2, Z-VLR-AMC was most preferred by FP-3 and BP-2 and boc-VLK-AMC was relatively well catalysed by all three. Substrates common between two of the three enzymes included Z-LQ-AMC, utilised by FP-2 and FP-3 only; whereas Bz-FVR-AMC, and surprisingly, Ac-DEVD-AMC were utilised by FP-3 and BP-2. In line with its less stringent substrate preference, FP-2 utilised substrates such as Z-FR-AMC, D-VLK-AMC and to a lesser extent Z-LE-AMC, PFR-AMC and Ac-FR-AMC, which were not hydrolysed by FP-3 or BP-2. Z-GPR-AMC was the only substrate unique to FP-3 and only Ac-NlePNleD-AMC was unique to BP-2; although the rate of catalysis in this latter case was consistently but reproducibly very low (~0.8% of the rate observed with Z-VLR-AMC). As previously reported [14], we found that BP-2 cleaved the tri-peptide substrate Z-VLR-AMC significantly more efficiently than di-peptide substrates (*e.g.* Z-LR-AMC). Since the residue positions were fixed relative to the cleavage site for the coumarin-based substrates, this precluded an understanding of 'prime-side' binding information. In this respect, we additionally sought a screening method providing synergistic residue interplay and which also spanned the active site. The use of FRET-based peptide substrates [20] was ideal for this purpose.

FRET-peptide based substrate screen

We synthesised a 1536 member FRET peptide substrate library (Supplementary Figure 1) which was screened against all three peptidases. In order to better present the wealth of data generated, the rates of fluorescence generation were converted to a colour transition intensity plot from a maximum rate, shown as yellow, to zero rate, shown as black (Figure 2). The data, presented as the library was synthesised, comprised sixteen 96-well plates. Each peptidase data set was processed independently of the others with the maximum rate observed for each peptidase used

as the maximum intensity. A simple visual assessment clearly shows that FP-2 displayed much greater diversity in its substrate preference compared with FP-3, whereas BP-2 displayed more restricted substrate preferences as indicated by significantly fewer peptides that displayed substrate activity. In order to avoid an in depth and lengthy description of the data, a summary of the salient features has been presented with the obvious caveat that synergistic exceptions may be present in the data (Supplementary Table 1). As a general principle, the presence of particular amino acids (*e.g.* Leu) at a given position, allowed greater residue tolerance at other positions; whereas the presence of other amino acids (*e.g*. Pro or Val) obliterated substrate activity. Based on the screening data, some of the highest ranking FRET peptide substrates were then selected for kinetic and structural analysis (Table 1). Also, the FRET substrate library did not contain a Leu-Arg containing sequence, and mindful of the peptide-AMC data (Figure 1), a representative set of Leu-Arg containing FRET peptides were additionally synthesised. The results clearly indicated that the Leu-Arg containing peptides were much better substrates, with second order rate constants generally ten-fold higher as a result of a reduction in K_m as well as an increase in k_{cat} (Table 1). The absence of Leu-Arg containing peptides in the substrate library may have serendipitously had a benefit, since the cleavage rates for these peptides would have dominated the data so as to mask the more subtle information. FP-2 exhibited a clear preference for arginine and glutamic acid as P1 residue in combination with leucine and to a lesser extent tyrosine as the P2 residue. The β-branched isoleucine was moderately well accepted as the P2 residue whereas the acidic aspartic acid clearly was not. Within this Leu/(Tyr)-Arg/Glu-P1'-P2' motif many combinations of P1'-P2' were accommodated with the exception of P1' proline. FP-3 exhibited a clear preference for arginine/lysine as the P1 residue and to a lesser extent 2-thienylalanine, threonine and glutamine in combination with leucine and to an equal extent tyrosine as the P2 residue. Clearly not accepted were the β-branched residue valine nor the acidic residue glutamic acid as the P2 residue. Within this P3-Leu/Tyr-Arg/Lys/(2-Thi)/(Thr)/(Gln)-P1' motif, combinations containing P1' proline, aspartic acid and isoleucine were not accommodated. BP-2 exhibited a clear preference for arginine/lysine and glycine as P1 residue and occasionally the aromatic 2-thienylalanine residue. The P1 arginine/lysine occurs in combination with predominantly leucine as the P2 residue but in stark contrast to FP-3, the β-branched

valine was also accommodated. The P1 glycine occurred in combination with predominantly the β-branched threonine or aromatic Thi as the P2 residue and clearly not accepted in this particular motif were acidic or larger aromatics. This preference for P1 glycine is very unusual for papain family peptidases. However a similar result was observed for vinckepain-2 through a positional scanning tetra-peptide AMC library [27].

Inhibitor profile

To extend the peptide substrate analysis, we tested a number of literature-based inhibitors against FP-2, FP-3, BP-2 and a range of human cathepsin peptidases. The inhibitors were split as two class types based on whether they were reversible or irreversible inhibitors (Tables 2 and 3). As a generalisation, the results reflected aspects of the substrate screening data where certain binding elements (*e.g.* Leu at P2) were required for recognition and other elements added to selectivity. As with the substrate results, the inhibitor data clearly highlighted differences between the inhibition of FP-2, FP-3 and BP-2 by the various compounds analysed. For example nitrile **1** [22] inhibited FP-2 and BP-2 in approximately the 500 nM range whereas FP-3 was not significantly inhibited at 2000 nM. Azepanone **4** [23] was a relatively potent inhibitor of FP-2 compared to FP-3 and BP-2 although the diacylcarbohydrazide **5** [36] was even more selective for FP-2 compared to FP-3 and BP-2. The azapeptides **6** and **7** [25] served to highlight the dramatic changes in potency and selectivity due to relatively minor structural changes. Whilst the aldehyde leupeptin **8** was a potent inhibitor of all three peptidases. The vinyl sulfone **12** (LHVS) [26], which has been utilised as a probe compound to correlate falcipain inhibition with antimalarial effect [37, 38], clearly demonstrated good potency against FP-2, FP-3 and BP-2, but when tested against a number of mammalian cathepsin peptidases it was non-selective (*e.g.* K_i values versus cathepsin B, S, K, L and V were 314.8 ± 19.9 nM, 0.21 ± 0.05 nM, 1.1 ± 0.1 nM, 10 ± 1 nM, 0.08 ± 0.03 nM respectively). Although in general the augmentation of compounds with potent electrophilic centres resulted in increased potency, these changes tended to reduce specificity as judged by the inhibition profile against the common human cathepsin peptidases. As an additional analysis of inhibition profile we tested a range of irreversible inhibitors against all three peptidases. The results of the rates of inhibition

and relative potencies of the irreversible inhibitors against FP-2, FP-3 and BP-2 were even more marked than those observed with the reversible inhibitors. For example E-64 (**9**) and E-64c (**10**), were relatively fast potent inhibitors of BP-2 and relatively slow poor inhibitors of FP-3 and FP-2 (Table 3). Conversely, the halomethylketones (**11**, **13-16**) (examples of which have been used to correlate the inhibition of falcipains with antimalarial activity [39, 40]) were fast potent inhibitors of FP-2 and FP-3 and relatively slow poor inhibitors of BP-2.

DISCUSSION

The papain-like cysteine peptidases FP-2 and FP-3 have been shown to play critical roles in haemoglobin hydrolysis as an essential requirement for parasite growth and development. As such they have been investigated as possible therapeutic targets for the treatment of malaria. In order to facilitate a medicinal chemistry programme, we analysed the substrate and inhibitor profiles of the two essential *P. falciparum* cysteine peptidases and, since the progression of compounds requires demonstration of efficacy in a disease-related animal model, a rodent ortholog of FP-2, *i.e.* BP-2, was also investigated.

The screen of AMC-based peptide substrates confirmed previously observed results and we were able to significantly extend the range of substrates analysed. We confirmed that Z-LR-AMC was most efficiently cleaved by FP-2 whereas Z-VLR-AMC most efficiently cleaved by FP-3 [6, 8]. The data obtained for BP-2 also agreed with previously reported results [14], even though this study had only analysed two substrates. New peptide-AMC substrates for all three peptidases were also identified. In this case the data obtained with the AMC peptide substrate complemented the FRET data and provided valuable information relating to sequences absent from the FRET library. The peptide-AMC substrates also proved to be excellent probes with which to study the effects of the N-terminal capping group on substrate catalysis and specificity. The FRET peptide substrate data clearly demonstrated that FP-2 utilised a greater diversity of substrates than FP-3, and BP-2 was relatively specific in its substrate preference. In both cases, FP-2 was shown to have significantly higher specific activity against the substrates tested than FP-3. Although a similar observation has been previously reported [8], it has been noted that on a molar basis, FP-3 was a more efficient haemoglobinase than FP-2 [8]. It is therefore possible that the *in vitro* assay conditions artificially favoured FP-2 peptidase activity, whereas in

the natural context these peptidases exhibit equipotent activity against protein substrates, where activity may be affected by the cellular environment [8]. Another consideration is that although FP-2 has been shown to be an efficient haemoglobinase, it is also capable of effectively cleaving erythrocyte membrane ankyrin and protein 4.1 [41]. There is therefore biochemical evidence to suggest that FP-2 can accommodate a wider selection of biological substrates, perhaps augmenting its primary role by assisting in other cellular functions. Recent evidence does however strongly suggest the primary role of FP-2 is the digestion of haemoglobin [42]. The homology models of FP-2 and FP-3 [43-45], as well as the structure of FP-2 [42] have been reported; however the absence of publicly available crystallographic data prevented a structure-based interpretation of the results. Although it was clear from amino acid sequence alignments that FP-2 and FP-3 were significantly different from BP-2 (data not shown), this data did not greatly facilitate our interpretations without gross assumptions. Irrespective of such interpretations, the data clearly indicated that the FP-2 substrate profile was dominated by Leu in the P2 position and the presence of Glu or especially Arg in P1 greatly enhanced the cleavage rates. FP-2 therefore had a specific binding preference for Arg in P1 as opposed to just being able to accommodate Arg in this position and also showed a selective preference for glutamic acid when compared with FP-3 and BP-2.

In contrast, the FP-3 substrate preference indicated a more extended substrate binding pocket with a preference for beta-branched or hydrophobic residues in P3. In this case the substitution of the P1 residue with Arg had a relatively modest effect on the rate of catalysis suggesting that other factors augmented substrate binding. Perhaps in this case backbone interactions provided a greater contribution to substrate binding. The FRET peptide substrates most efficiently cleaved by BP-2 were of two distinct classes: one comprised capped tri-peptides and the other comprised capped tetra-peptides with, in the latter cases, an absolute preference for Gly in P1. Once again this suggested an extended binding mode for BP-2 substrates with at least two types of mode being observed. Again perhaps backbone interactions contributed significantly to substrate binding and recognition.

The characterisation of the substrate and inhibitor profiles of FP-2, FP-3 and BP-2 will significantly aid the biochemical understanding of the enzymes and also facilitate aspects of the structure-based drug discovery process. Our results emphasise the complexity of parasite biology highlighting not only subtle differences between

highly similar enzymes from the same species but also significant differences between similar enzymes from different species. The malaria cysteine peptidase drug discovery process has therefore to take into account the inhibition of more than one target. Clearly for *P. falciparum*, FP-2 and FP-3 will need to be inhibited whereas for *P. berghei*, BP-2 will need to be inhibited. Since we and others have obtained data demonstrating clear and marked differences between these enzymes, consideration has to be given to the idea that a single compound with potent inhibition of FP-2 and FP-3, other than those that contain a potent electrophilic centre, may be difficult to achieve; whilst inhibition of all three peptidases will likely prove even more challenging. Another level of complication is the variations one may observe for an enzyme from different parasite strains. Coupled to this, the data would suggest that an inhibitor of FP-2 and/or FP-3 may demonstrate efficacy in a *P. falciparum* cell-based assay and yet possibly show no efficacy in a *P. berghei* disease related animal model. Clearly our results, as well as other reports, indicate that a degree of caution must be exercised when interpreting data generated from the *P. berghei* disease related animal model when testing FP-2 and FP-3 inhibitors. It has been proposed that BP-2 does not fulfil the battery of roles as does FP-2 [14] and it has been suggested that the *P. berghei* disease related animal model may not represent an adequately predictive model for the *in vivo* analysis of *P. falciparum* inhibitors [14, 46]. Although the current animal mouse models provide a framework for the testing and development of anti-malarial therapeutics [17], continued effort should be directed towards the development of more pertinent models (*e.g.* transgenic or immunocompromised models [18]). Given the current restrictions on the types of cell-based assays and disease related animal models readily available for malaria research, care should be exercised when evaluating the efficacy of compounds, whether they are cysteine protease inhibitors or even other therapeutics targets. In the worst case scenario, it is possible that a compound aimed at treating *P. falciparum* infection in the human host is rejected on the basis that it shows no efficacy in a *P. berghei* disease related mouse animal model.

ACKNOWLEDGMENTS

We thank the staff at Amura Therapeutics Limited for advice and manuscript review. We are indebted to Prof. P.J. Rosenthal for provision of malarial peptidases, manuscript review, advice and support.

REFERENCES

1 (2002) Nature insight- Malaria. In Nature (Weiss, U., ed.), pp. 669-715, Macmillan Magazines Ltd., London

2 Bozdech, Z., Llinas, M., Pulliam, B. L., Wong, E. D., Zhu, J. and DeRisi, J. L. (2003) The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. PLoS Biol **1**, E5

3 Rosenthal, P. J. (2001) Protease inhibitors. In Antimalarial Chemotherapy. Mechanisms of action, resistance, and new directions in drug discovery (Rosenthal, P. J., ed.), Humana Press, Totowa, NJ, USA.

4 Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T., James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., Craig, A., Kyes, S., Chan, M. S., Nene, V., Shallom, S. J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M. W., Vaidya, A. B., Martin, D. M., Fairlamb, A. H., Fraunholz, M. J., Roos, D. S., Ralph, S. A., McFadden, G. I., Cummings, L. M., Subramanian, G. M., Mungall, C., Venter, J. C., Carucci, D. J., Hoffman, S. L., Newbold, C., Davis, R. W., Fraser, C. M. and Barrell, B. (2002) Genome sequence of the human malaria parasite Plasmodium falciparum. Nature **419**, 498-511

5 Sijwali, P. S., Kato, K., Seydel, K. B., Gut, J., Lehman, J., Klemba, M., Goldberg, D. E., Miller, L. H. and Rosenthal, P. J. (2004) Plasmodium falciparum cysteine protease falcipain-1 is not essential in erythrocytic stage malaria parasites. Proc Natl Acad Sci U S A **101**, 8721-8726

6 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

7 Singh, N., Sijwali, P. S., Pandey, K. C. and Rosenthal, P. J. (2006) Plasmodium falciparum: Biochemical characterization of the cysteine protease falcipain-2'. Exp Parasitol **112**, 187-192

8 Sijwali, P. S., Shenai, B. R., Gut, J., Singh, A. and Rosenthal, P. J. (2001) Expression and characterization of the Plasmodium falciparum haemoglobinase falcipain-3. Biochem J **360**, 481-489

9 Greenbaum, D. C., Baruch, A., Grainger, M., Bozdech, Z., Medzihradszky, K. F., Engel, J., DeRisi, J., Holder, A. A. and Bogyo, M. (2002) A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. Science **298**, 2002-2006

10 Eksi, S., Czesny, B., Greenbaum, D. C., Bogyo, M. and Williamson, K. C. (2004) Targeted disruption of Plasmodium falciparum cysteine protease, falcipain 1, reduces oocyst production, not erythrocytic stage growth. Mol Microbiol **53**, 243-250

11 Malhotra, P., Dasaradhi, P. V., Kumar, A., Mohmmed, A., Agrawal, N., Bhatnagar, R. K. and Chauhan, V. S. (2002) Double-stranded RNA-mediated gene silencing of cysteine proteases (falcipain-1 and -2) of Plasmodium falciparum. Mol Microbiol **45**, 1245-1254

12 Singh, N., Sijwali, P. S., Pandey, K. C. and Rosenthal, P. J. (2005) Plasmodium falciparum: Biochemical characterization of the cysteine protease falcipain-2'. Exp Parasitol

13 Sim, T. S., Loke, P., Lee, M. A., Singh, M. and Flotow, H. (2001) Cloning and sequence characterisation of falcipain-2 from Plasmodium falciparum Gombak A strain (Malaysia). Parasitol Res **87**, 683-686

14 Chan, C., Goh, L. L. and Sim, T. S. (2005) Differences in biochemical properties of the Plasmodial falcipain-2 and berghepain-2 orthologues: implications for in vivo screens of inhibitors. FEMS Microbiol Lett **249**, 315-321

15 Sijwali, P. S. and Rosenthal, P. J. (2004) Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by Plasmodium falciparum. Proc Natl Acad Sci U S A **101**, 4384-4389

16 Dahl, E. L. and Rosenthal, P. J. (2005) Biosynthesis, localization, and processing of falcipain cysteine proteases of Plasmodium falciparum. Mol Biochem Parasitol **139**, 205-212

17 Fidock, D. A., Rosenthal, P. J., Croft, S. L., Brun, R. and Nwaka, S. (2004) Antimalarial drug discovery: efficacy models for compound screening. Nat Rev Drug Discov **3**, 509-520

18 Moreno, A., Badell, E., Van Rooijen, N. and Druilhe, P. (2001) Human malaria in immunocompromised mice: new in vivo model for chemotherapy studies. Antimicrob Agents Chemother **45**, 1847-1853

19 Atherton, E. and Sheppard, R. C. (1989) Solid phase peptide synthesis: A practical approach. IRL Press, Oxford

20 Meldal, M. and Breddam, K. (1991) Anthranilamide and nitrotyrosine as a donor-acceptor pair in internally quenched fluorescent substrates for endopeptidases: multicolumn peptide synthesis of enzyme substrates for subtilisin Carlsberg and pepsin. Anal Biochem **195**, 141-147

21 Grabowska, U., Rizzo, A., Farnell, K. and Quibell, M. (2000) 5-(hydroxymethyl)oxazoles: versatile scaffolds for combinatorial solid-phase synthesis of 5-substituted oxazoles. J Comb Chem **2**, 475-490

22 Missbach, M. (2001) Dipeptide nitrile cathepsin K inhibitors. In World Intellectual Property Organisation (Organisation, W. I. P., ed.), Novartis AG, Swizterland

23 Marquis, R. W., Ru, Y., LoCastro, S. M., Zeng, J., Yamashita, D. S., Oh, H. J., Erhard, K. F., Davis, L. D., Tomaszek, T. A., Tew, D., Salyers, K., Proksch, J., Ward, K., Smith, B., Levy, M., Cummings, M. D., Haltiwanger, R. C., Trescher, G., Wang, B., Hemling, M. E., Quinn, C. J., Cheng, H. Y., Lin, F., Smith, W. W., Janson, C. A., Zhao, B., McQueney, M. S., D'Alessio, K., Lee, C. P., Marzulli, A., Dodds, R. A., Blake, S., Hwang, S. M., James, I. E., Gress, C. J., Bradley, B. R., Lark, M. W., Gowen, M. and Veber, D. F. (2001) Azepanone-based inhibitors of human and rat cathepsin K. J Med Chem **44**, 1380-1395

24 Thompson, S. K., Halbert, S. M., Bossard, M. J., Tomaszek, T. A., Levy, M. A., Zhao, B., Smith, W. W., Abdel-Meguid, S. S., Janson, C. A., D'Alessio, K. J., McQueney, M. S., Amegadzie, B. Y., Hanning, C. R., DesJarlais, R. L., Briand, J., Sarkar, S. K., Huddleston, M. J., Ijames, C. F., Carr, S. A., Garnes, K. T., Shu, A., Heys, J. R., Bradbeer, J., Zembryki, D., Lee-Rykaczewski, L., James, I. E., Lark, M. W., Drake, F. H., Gowen, M., Gleason, J. G. and Veber, D. F. (1997) Design of potent and selective human cathepsin K inhibitors that span the active site. Proc Natl Acad Sci U S A **94**, 14249-14254

25 Xing, R. and Hanzlik, R. P. (1998) Azapeptides as inhibitors and active site titrants for cysteine proteinases. J Med Chem **41**, 1344-1351

26 Palmer, J. T., Rasnick, D., Klaus, J. L. and Bromme, D. (1995) Vinyl sulfones as mechanism-based cysteine protease inhibitors. J Med Chem **38**, 3193-3196

27 Singh, A., Shenai, B. R., Choe, Y., Gut, J., Sijwali, P. S., Craik, C. S. and Rosenthal, P. J. (2002) Critical role of amino acid 23 in mediating activity and specificity of vinckepain-2, a papain-family cysteine protease of rodent malaria parasites. Biochem J **368**, 273-281

28 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

29 Barrett, A. J., Rawlings, N. D. and Woessner, J. F. (2000) Handbook of proteolytic enzymes. Academic Press, London

30 Cornish-Bowden, A. (2001) Fundementals of enzyme kinetics. Portland Press, London

31 Stone, S. R. and Hofsteenge, J. (1986) Kinetics of the inhibition of thrombin by hirudin. Biochemistry **25**, 4622-4628

32 Morrison, J. F. and Walsh, C. T. (1988) The behavior and significance of slow-binding enzyme inhibitors. Adv Enzymol Relat Areas Mol Biol **61**, 201-301

33 Cleland, W. W. (1964) Dithiothreitol, a New Protective Reagent for Sh Groups. Biochemistry **3**, 480-482

34 Maly, D. J., Leonetti, F., Backes, B. J., Dauber, D. S., Harris, J. L., Craik, C. S. and Ellman, J. A. (2002) Expedient solid-phase synthesis of fluorogenic protease substrates using the 7-amino-4-carbamoylmethylcoumarin (ACC) fluorophore. J Org Chem **67**, 910-915

35 Na, B. K., Shenai, B. R., Sijwali, P. S., Choe, Y., Pandey, K. C., Singh, A., Craik, C. S. and Rosenthal, P. J. (2004) Identification and biochemical characterization of vivapains, cysteine proteases of the malaria parasite Plasmodium vivax. Biochem J **378**, 529-538

36 Bossard, M. J., Tomaszek, T. A., Levy, M. A., Ijames, C. F., Huddleston, M. J., Briand, J., Thompson, S., Halpert, S., Veber, D. F., Carr, S. A., Meek, T. D. and Tew, D. G. (1999) Mechanism of inhibition of cathepsin K by potent, selective 1, 5-diacylcarbohydrazides: a new class of mechanism-based inhibitors of thiol proteases. Biochemistry **38**, 15893-15902

37 Rosenthal, P. J., Olson, J. E., Lee, G. K., Palmer, J. T., Klaus, J. L. and Rasnick, D. (1996) Antimalarial effects of vinyl sulfone cysteine proteinase inhibitors. Antimicrob Agents Chemother **40**, 1600-1603

38 Shenai, B. R., Lee, B. J., Alvarez-Hernandez, A., Chong, P. Y., Emal, C. D., Neitz, R. J., Roush, W. R. and Rosenthal, P. J. (2003) Structure-activity relationships for inhibition of cysteine protease activity and development of Plasmodium falciparum by peptidyl vinyl sulfones. Antimicrob Agents Chemother **47**, 154-160

39 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

40 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

41 Dua, M., Raphael, P., Sijwali, P. S., Rosenthal, P. J. and Hanspal, M. (2001) Recombinant falcipain-2 cleaves erythrocyte membrane ankyrin and protein 4.1. Mol Biochem Parasitol **116**, 95-99

42 Pandey, K. C., Wang, S. X., Sijwali, P. S., Lau, A. L., McKerrow, J. H. and Rosenthal, P. J. (2005) The Plasmodium falciparum cysteine protease falcipain-2 captures its substrate, hemoglobin, via a unique motif. Proc Natl Acad Sci U S A **102**, 9138-9143

43 Sabnis, Y., Rosenthal, P. J., Desai, P. and Avery, M. A. (2002) Homology modeling of falcipain-2: validation, de novo ligand design and synthesis of novel inhibitors. J Biomol Struct Dyn **19**, 765-774

44 Sabnis, Y. A., Desai, P. V., Rosenthal, P. J. and Avery, M. A. (2003) Probing the structure of falcipain-3, a cysteine protease from Plasmodium falciparum: comparative protein modeling and docking studies. Protein Sci **12**, 501-509

45 Goh, L. L. and Sim, T. S. (2004) Homology modeling and mutagenesis analyses of Plasmodium falciparum falcipain 2A: implications for rational drug design. Biochem Biophys Res Commun **323**, 565-572

46 Humphreys, M. J., Moon, R. P., Klinder, A., Fowler, S. D., Rupp, K., Bur, D., Ridley, R. G. and Berry, C. (1999) The aspartic proteinase from the rodent parasite Plasmodium berghei as a potential model for plasmepsins from the human malaria parasite, Plasmodium falciparum. FEBS Lett **463**, 43-48

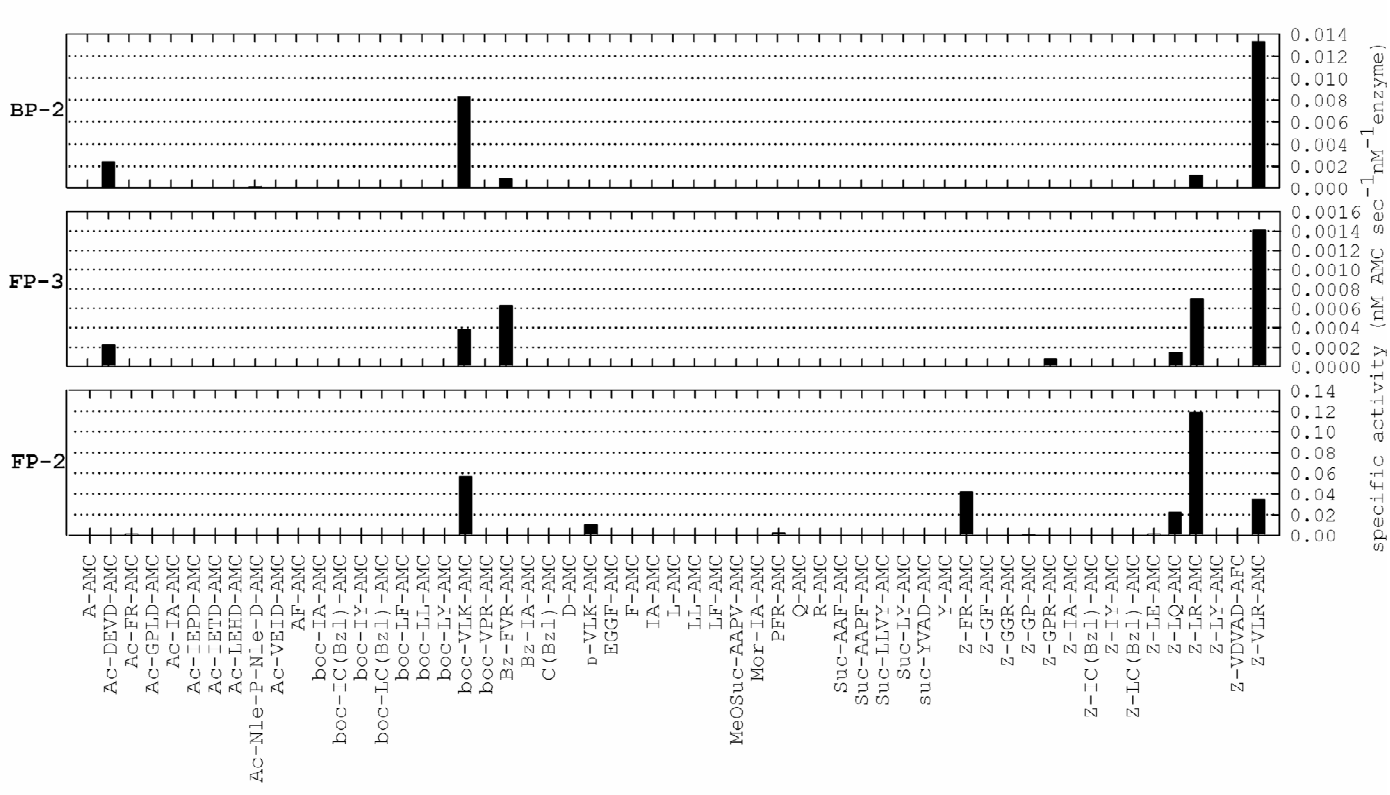


Figure 1 The utilisation of AMC-based peptide substrates by FP-2, FP-3 and BP-2

The rates of substrate catalysis were determined using fixed concentrations of peptide-AMC substrates (10 µM final) in the presence of either FP-2 (1 nM), FP-3 (10 nM) of BP-2 (5 nM).

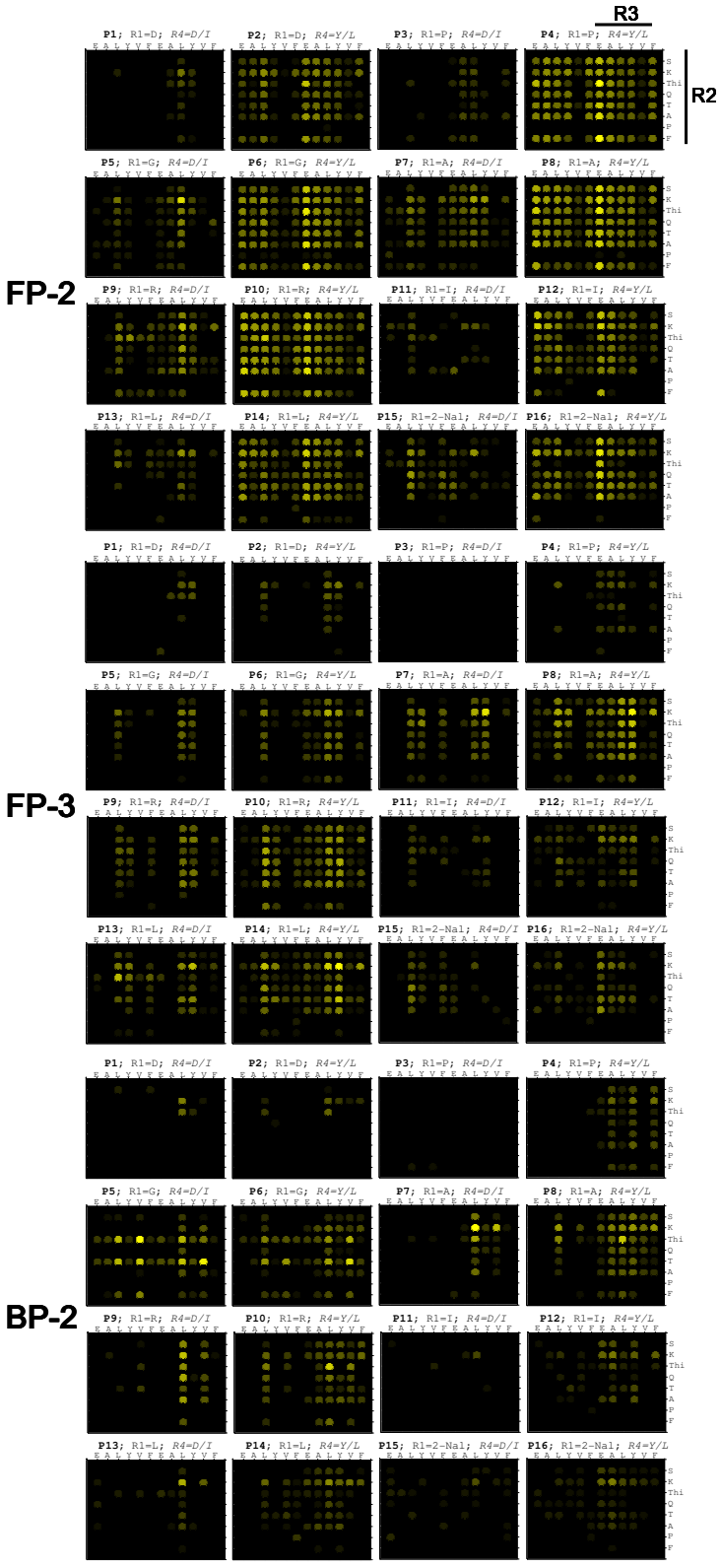


Figure 2 Graphical representation of the rates of FRET peptide substrate catalysis as intensity plots

The rates of fluorescence generation (*i.e.* RFU/sec) from the FRET peptide library screen (10 µM final), in the presence of either FP-2 (1 nM), FP-3 (10 nM) or BP-2 (5 nM), were independently converted into colour-transition plots for each data set. The library format, construction and key have been described in detail elsewhere (Supplementary Figure 1).

Biochemical Journal Immediate Publication. Published on 15 June 2006 as manuscript BJ20060422

	Sequence	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
FP-2				
	Abz-**Leu**-Arg▼*Phe*-Pro-Tyr(3-NO_2)-Glu-NH_2	0.79 ± 0.01	0.9 ± 0.1	8.4×10^5
	Abz-**Leu**-Glu▼*Phe*-Pro-Tyr(3-NO_2)-Glu-NH_2	0.45 ± 0.01	7.7 ± 1.4	5.9×10^4
	Abz-**Leu**-Arg▼*Phe*-Gly-Tyr(3-NO_2)-Glu-NH_2	0.746 ± 0.002	1.2 ± 0.1	6.2×10^5
	Abz-**Leu**-Glu▼*Phe*-Gly-Tyr(3-NO_2)-Glu-NH_2	0.45 ± 0.01	9 ± 2	5.0×10^4
	Abz-**Leu**-Arg▼*Ala*-Gly-Tyr(3-NO_2)-Glu-NH_2	0.645 ± 0.004	1.5 ± 0.1	4.3×10^5
	Abz-**Leu**-Glu▼*Ala*-Gly-Tyr(3-NO_2)-Glu-NH_2	0.396 ± 0.012	11 ± 3	3.6×10^4
	Abz-**Leu**-Arg▼*Ala*-Pro-Tyr(3-NO_2)-Glu-NH_2	0.622 ± 0.005	2 ± 0.2	3.3×10^5
	Abz-**Leu**-Glu▼*Phe*-Ala-Tyr(3-NO_2)-Glu-NH_2	0.458 ± 0.009	4.5 ± 1	1.0×10^5
	Abz-**Leu**-Glu▼*Thi*-Pro-Tyr(3-NO_2)-Glu-NH_2	0.40 ± 0.01	4 ± 1	9.9×10^4
	Abz-**Leu**-Glu▼*Ser*-Ala-Tyr(3-NO_2)-Glu-NH_2	0.40 ± 0.01	6 ± 1.5	6.5×10^4
	Abz-**Leu**-Glu▼*Lys*-Ile-Tyr(3-NO_2)-Glu-NH_2	0.42 ± 0.01	7 ± 1	5.8×10^4
	Abz-**Leu**-Glu▼*Thi*-Ala-Tyr(3-NO_2)-Glu-NH_2	0.52 ± 0.01	9 ± 1.6	5.8×10^4
FP-3				
	Abz-Leu-**Leu**-Arg▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.204 ± 0.005	4 ± 1	5.5×10^4
	Abz-Leu-**Leu**-Lys▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.0389 ± 0.0002	2.7 ± 0.2	1.4×10^4
	Abz-Leu-**Tyr**-Arg▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.744 ± 0.002	20 ± 3	3.8×10^4
	Abz-Leu-**Tyr**-Lys▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.106 ± 0.001	8 ± 0.8	1.3×10^4
	Abz-Leu-**Leu**-Lys▼*Leu*-Tyr(3-NO_2)-Glu-NH_2	0.049 ± 0.001	5 ± 1	9.7×10^3
	Abz-Leu-**Tyr**-Lys▼*Leu*-Tyr(3-NO_2)-Glu-NH_2	0.139 ± 0.003	15 ± 2	9.1×10^3
	Abz-Tyr-**Leu**-Lys▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.0477 ± 0.0004	6 ± 0.4	8.4×10^3
	Abz-Leu-**Tyr**-Thr▼*Leu*-Tyr(3-NO_2)-Glu-NH_2	0.037 ± 0.001	4.5 ± 1	8.3×10^3
	Abz-Leu-**Tyr**-Thi▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.036 ± 0.001	5 ± 1	7.7×10^3
	Abz-Leu-**Leu**-Lys▼*Gly*-Tyr(3-NO_2)-Glu-NH_2	0.052 ± 0.001	7 ± 1	7.7×10^3
	Abz-Leu-**Tyr**-Thr▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.0416 ± 0.0003	6 ± 0.4	7.2×10^3
	Abz-Ile-**Leu**-Lys▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.083 ± 0.001	13 ± 2	6.3×10^3
BP-2				
	Abz-Leu-**Leu**-Arg▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.0313 ± 0.0004	0.5 ± 0.1	6.1×10^4
	Abz-Ile-**Leu**-Lys▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.129 ± 0.002	2.6 ± 0.6	5.0×10^4
	Abz-Leu-**Tyr**-Arg▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.074 ± 0.001	1.8 ± 0.2	4.1×10^4
	Abz-Leu-**Leu**-Thi▼*Ala*-Tyr(3-NO_2)-Glu-NH_2	0.099 ± 0.001	3 ± 0.5	3.3×10^4
	Abz-Leu-**Leu**-Thi▼*Arg*-Tyr(3-NO_2)-Glu-NH_2	0.131 ± 0.002	4 ± 1	3.1×10^4
	Abz-Ile-**Leu**-Lys▼*Leu*-Tyr(3-NO_2)-Glu-NH_2	0.128 ± 0.002	7 ± 1	1.8×10^4
	Abz-Ile-**Leu**-Lys▼*Arg*-Tyr(3-NO_2)-Glu-NH_2	0.143 ± 0.003	8 ± 2	1.7×10^4
	Abz-**Leu**-Arg▼*Phe*-Pro-Tyr(3-NO_2)-Glu-NH_2	0.189 ± 0.003	12 ± 2	1.6×10^4
	Abz-Ile-Val-**Thr**-Gly▼*Tyr(3-NO_2)*-Glu-NH_2	0.465 ± 0.008	30 ± 3	1.5×10^4
	Abz-Ile-Leu-**Thi**-Gly▼*Tyr(3-NO_2)*-Glu-NH_2	0.201 ± 0.003	15 ± 2	1.3×10^4
	Abz-Asp-Val-**Thr**-Gly▼*Tyr(3-NO_2)*-Glu-NH_2	0.39 ± 0.02	31 ± 7	1.3×10^4
	Abz-Ile-Leu-**Thr**-Gly▼*Tyr(3-NO_2)*-Glu-NH_2	0.33 ± 0.01	27 ± 4	1.2×10^4
	Abz-Asp-Val-**Thi**-Gly▼*Tyr(3-NO_2)*-Glu-NH_2	0.91 ± 0.06	75 ± 21	1.2×10^4
	Abz-Leu-Val-**Thi**-Gly▼*Tyr(3-NO_2)*-Glu-NH_2	0.71 ± 0.03	61 ± 12	1.2×10^4

Table 1 Kinetic and structural characterisation selected FRET peptide substrates

The P2 (bold), P1 (white on black background) and P1' (italics) positions have been highlighted relative to the cleavage site (▼).

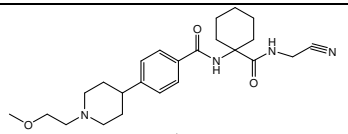
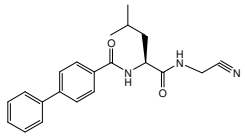
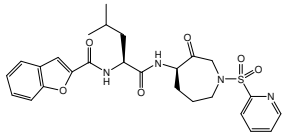
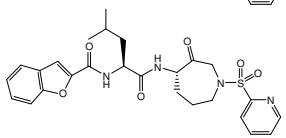
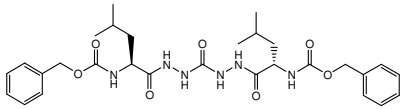
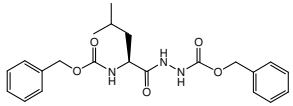
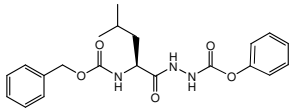
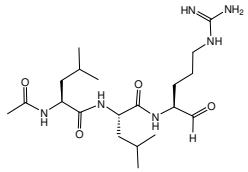
Compound	Structure	K_i (nM)							
		FP-2	FP-3	BP-2	cat B	cat S	cat L	cat K	cat V
1		676.8 ± 71.1	>2000	414.8 ± 55.0	>2000	>2000	1790 ± 240	5.5 ± 0.4	>2000
2		181 ± 18	943 ± 135	644 ± 55	>2000	582 ± 157	262 ± 30	3.6 ± 0.6	57 ± 7
3		>2000	>2000	>2000	>2000	>2000	>2000	285 ± 34	>2000
4		283.3 ± 27.5	1164 ± 154	>2000	>2000 (500)[23]	26.8 ± 2.1 (4)[23]	240 ± 46 (2.2)[23]	1.8 ± 0.2 (0.16)[23]	19.4 ± 2.4
5		31.8 ± 4.1	728 ± 111	419 ± 51	2311 ± 324	2185 ± 796	188 ± 16	8.4 ± 1.3 (2.7 ± 0.2) [36]	0.28 ± 0.04
6		>2000	>2000	378 ± 51	>2000	179 ± 70	976 ± 130	159 ± 10	20.3 ± 3.1
7		141 ± 17	1364 ± 170	43.4 ± 5.4	934 ± 82	26.8 ± 6.7	30.3 ± 2.6	7.7 ± 0.5	0.58 ± 0.12
8 (Leupeptin)		31.9 ± 4.0	98 ± 12	19.1 ± 2.4	47.1 ± 2.4	20.2 ± 1.7	2.6 ± 0.2	20.4 ± 3.2	11.8 ± 1.3

Table 2 The inhibition constants of selected reversible compounds against FP-2, FP-3, BP-2 and selected human cathepsin peptidases

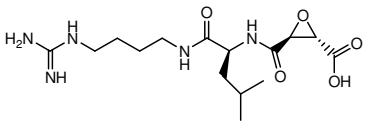
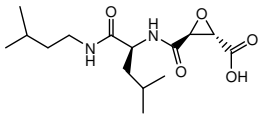
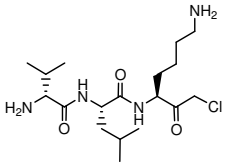
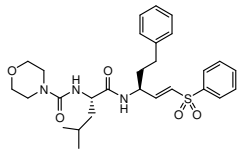
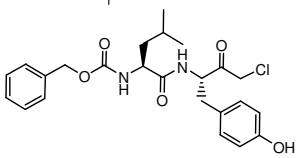
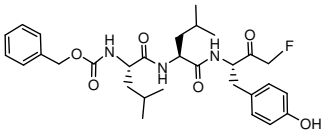
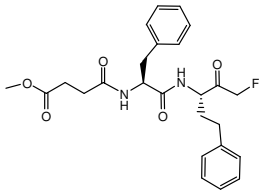
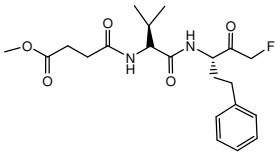
Name	Structure	k_{inact} ($M^{-1}s^{-1}$)		
		FP-2	FP-3	BP-2
9 (E-64)		$1.16 \pm 0.01 \times 10^3$	$6.6 \pm 0.3 \times 10^2$	$> 1 \times 10^7$
10 (E-64c)		$7.6 \pm 0.6 \times 10^2$	$1.0 \pm 0.1 \times 10^2$	$1.8 \pm 0.2 \times 10^5$
11 (D-VLK-CMK)		$1.6 \pm 0.2 \times 10^6$	$5.8 \pm 1.9 \times 10^4$	$5.3 \pm 0.2 \times 10^3$
12 (LHVS)		$1.2 \pm 0.2 \times 10^4$ (7.86×10^4) [38]	$1.4 \pm 0.4 \times 10^3$ (5.71×10^4) [38]	$9.7 \pm 0.8 \times 10^3$
13 (Z-LY-CMK)		$5.7 \pm 0.2 \times 10^5$	$> 1 \times 10^7$	$9.1 \pm 0.2 \times 10^3$
14 (Z-LLY-FMK)		$> 1 \times 10^7$	$> 1 \times 10^7$	$3.5 \pm 1.1 \times 10^3$
15 (Mu-Phe-hPhe-FMK)		$> 1 \times 10^7$	$> 1 \times 10^7$	$<1 \times 10^2$
16 (Mu-Val-hPhe-FMK)		$> 1 \times 10^7$	$> 1 \times 10^7$	$7.8 \pm 1.1 \times 10^2$

Table 3 Second order inactivation rate constants for selected irreversible compounds against FP-2, FP-3, BP-2

Synthesis of Z-Leu-Arg-AMC (Z-LR-AMC)

H-Arg-AMC (250 mg, 754.4 µmol), suspended in DMF:methanol (20:20 ml) containing NMM (166 µl; 1508.5 µmol), was stirred at room temperature for 150 min. with Z-Leu-O-succinimidyl ester (Bachem) (1360 mg, 3772 µmol). The reaction mixture was reduced *in vacuo* and triturated three times with toluene (30 ml) to leave a crude white gum. The crude mixture was dissolved in DCM:methanol (30:3 ml), Flash Si II (3.5 g; Argonaut Technologies, Hengoed, Wales, U.K.) added and the mixture reduced *in vacuo*. The dried material was dry-packed onto a FLASH Si II column (20 g; Argonaut) and subsequently eluted with a stepwise gradient consisting of 50 ml of each of the following- 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 8%, 10%, 12%, 16%, 18%, 20% and 25% methanol in DCM. Product-containing fractions were pooled and reduced *in vacuo* to leave a gum which was triturated three times with ethyl acetate:heptane (30 ml) to give Z-Leu-Arg-AMC as a white solid, yield: 356.7 mg (616 µmol, 82 %), ESI-MS 579.3 $[M + H]^+$ (calc. 578.66), with retention time 15.7 min. (>95%).

Synthesis of Z-Val-Leu-Arg-AMC (Z-VLR-AMC)

H-Arg-AMC (250 mg, 754.4 µmol), suspended in DMF:methanol (20:20 ml) containing NMM (166 µl; 1508.5 µmol), was stirred at room temperature for 195 min. with Boc-Leu-O-succinimidyl ester (Bachem) (1230 mg, 3772 µmol). The reaction mixture was reduced *in vacuo* and triturated three times with toluene (30 ml) to leave a crude gum. The crude mixture was dissolved in DCM:methanol (30:3 ml), Flash Si II (3.5 g; Argonaut Technologies) added and the mixture reduced *in vacuo*. The dried material was dry-packed onto a FLASH Si II column (20 g; Argonaut Technologies) and subsequently eluted with a stepwise gradient consisting of 50 ml of each of the following- 0%, 2%, 4%, 6%, 8%, 10%, 12%, 14%, 16%, 18%, 20% and 25% methanol in DCM. Product containing fractions were pooled and reduced *in vacuo* to leave a gum. The gum was suspended in 4M HCl in dioxane (30 ml; Aldrich) and the reaction stirred vigorously until the Boc de-protection reaction was complete as judged by thin-layer chromatography and LC-MS. The reaction was reduced *in vacuo* and triturated twice with diethylether (50 ml) to leave a crude solid. The solid was subsequently dissolved in DMF (30 ml) containing NMM (170 µl; 1532 µmol) and Z-Val-O-succinimidyl ester (1340 mg; 3830 µmol) added to the stirring reaction. After stirring overnight the reaction was reduced *in vacuo* and triturated three times with

toluene (30 ml) to leave a crude gum. The dried material was purified by flash chromatography as described above. Product containing fractions were pooled and reduced *in vacuo* to leave a gum which was triturated three times with ethyl acetate:heptane (30 ml) to give solid product: 276.3 mg (408 µmol; 53 %), ESI-MS 678.4 $[M + H]^+$ (calc. 677.35), with retention time 16.7 min. (>95%).