Naturally-occurring and recombinant forms of the aspartic proteinases plasmepsins I and II from the human malaria parasite *Plasmodium falciparum*

Lorraine Tyas[a], Ilya Gluzman[b], Richard P. Moon[c], Katharina Rupp[c], Jennifer Westling[d], Robert G. Ridley[c], John Kay[a], Daniel E. Goldberg[b], Colin Berry[a,*]

[a] *Cardiff School of Biosciences, Cardiff University, P.O. Box 911, Cardiff CF1 3US, UK*
[b] *Howard Hughes Medical Institute, Departments of Molecular Microbiology and Medicine, Washington University School of Medicine, St. Louis, MO 631110, USA*
[c] *F. Hoffmann-La Roche, Pharmaceuticals Division, Pharma Research Preclinical, CH-4070 Basel, Switzerland*
[d] *Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, FL 32610, USA*

Received 20 April 1999; received in revised form 7 June 1999

Abstract Comparable kinetic parameters were derived for the hydrolysis of peptide substrates and the interaction of synthetic inhibitors with recombinant and naturally-occurring forms of plasmepsin II. In contrast, recombinant plasmepsin I was extended by 12 residues at its N-terminus relative to its naturally-occurring counterpart and a 3–10-fold diminution in the k_{cat} values was measured for substrate hydrolysis by the recombinant protein. However, comparable K_i values were derived for the interaction of two distinct inhibitors with both forms of plasmepsin I, thereby validating the use of recombinant material for drug screening. The value of plasmepsin I inhibitors was determined by assessing their selectivity using human aspartic proteinases.

Key words: Malaria; Aspartic proteinase; Plasmepsin; *Plasmodium falciparum*

1. Introduction

Two aspartic proteinases (plasmepsin I and plasmepsin II) of the human malarial parasite *Plasmodium falciparum* play key roles in the essential pathway by which the parasites catabolise up to 80% of host cell haemoglobin during the intra-erythrocytic stages of their life-cycle [1,2]. Inhibition of the plasmepsins (by pepstatin) results in a cessation of parasite growth leading to death [3]. Thus the aspartic proteinases represent excellent targets for the design of new anti-malarial drugs.

Isolation of the naturally-occurring plasmepsins from the digestive vacuoles of the parasite in sufficient quantities to permit detailed investigations is impractical and therefore it has been necessary to produce each enzyme in recombinant form for initial characterisation [4,5]. Recombinant plasmepsin I appears to be more susceptible than recombinant plasmepsin II to several inhibitors (e.g. Ro40-4388 and SC-50083) which have been demonstrated to be effective in preventing haemoglobin degradation and the growth of *P. falciparum* in human red blood cells in culture [1,5]. Plasmepsin I was thus proposed as the primary drug target.

*Corresponding author. Fax: (+44) (1222) 874116.
E-mail: Berry@Cardiff.ac.uk

In order to validate whether recombinant forms of plasmepsins I and II were appropriate for use in systematic investigations into inhibitor design and development, it was considered vital to establish that the properties of the recombinant enzymes, in terms of activity and specificity, reflect those of their respective naturally-occurring counterparts.

2. Materials and methods

2.1. Preparation and N-terminal sequencing of recombinant forms of plasmepsins I and II

Genes encoding a truncated zymogen of plasmepsin II and an activation site mutant of the plasmepsin I zymogen capable of autoactivation [5] were expressed in *Escherichia coli* and the resultant semi-proplasmepsins I and II that accumulated were refolded and purified as described previously [4,5]. The autoactivation of each semi-proplasmepsin was carried out by mixing a sample of each purified preparation with one tenth the volume of 1 M sodium formate buffer pH 4.4 (1 M ionic strength), followed by incubation at 37°C for 30 min. The proteins that resulted from each autoactivation were resolved on (separate) 10% SDS-polyacrylamide gels and blotted onto polyvinylidene difluoride (PVDF) membranes. N-terminal sequencing of each protein band was performed by automated Edman degradation using an Applied Biosystems Protein Sequencer, model 476A (by courtesy of Dr. C.W. van den Berg, University of Wales College of Medicine, Cardiff, UK).

2.2. Purification of naturally-occurring plasmepsin I and plasmepsin II

Naturally-occurring plasmepsins I and II were isolated from the digestive vacuoles of *P. falciparum* as described previously [2,6]. Each of these enzyme preparations was obtained after hydroxylapatite chromatography and contained (respectively) the single proteolytic activity of either plasmepsin I or plasmepsin II without contamination by falcipain, the cysteine proteinase that is also present in the digestive vacuole [2].

2.3. Human aspartic proteinases

Naturally-occurring human liver cathepsin D (Athens Research and Technology Inc., Athens, GA, USA), human gastricsin and human pepsin were prepared as previously described [7] whilst recombinant human cathepsin E was purified from *E. coli* as described by Hill et al. [8].

2.4. Spectrophotometric analysis of aspartic proteinase activity

The hydrolysis of chromogenic substrates was monitored by following the decrease in absorbance at 300 nm using a Hewlett-Packard 8452A diode array spectrophotometer. Initial velocities were measured with at least six concentrations of each peptide substrate within an appropriate range in order to derive kinetic constants K_m and V_{max}. The estimated error for each parameter was always $< 15\%$. The concentration of active enzyme [E] used in each assay was determined by active site titration with isovaleryl pepstatin as described

PII: S0014-5793(99)00805-4

Table 1
Comparison of the kinetic parameters derived for the hydrolysis of chromogenic substrates by recombinant and naturally-occurring plasmepsins I and II

	P_5	P_4	P_3	P_2	$P_{1.5}$	$*P_1'$	P_2'	P_3'	Plasmepsin II						Plasmepsin I					
									Recombinant			Naturally occurring			Recombinant			Naturally occurring		
									K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
1	Leu	Glu	Arg	Ile	Phe	*Nph	Ser	Phe	10	13	1300	6	10	1670	10	2.1	210	8	16	2000
2	Lys	Glu	**Phe**	**Val**	Phe	*Nph	**Ala**	Leu	7	17	2430	7	19	2710	5	3.5	700	10	18	1800
3	Lys	Glu	**Leu**	Val	Phe	*Nph	Ala	Leu	5	10	2000	–	–	n.d.	6	1.5	250	8	6	750
4	Lys	Glu	Phe	**Asn**	Phe	*Nph	Ala	Leu	25	7	280	–	–	n.d.	20	2.0	100	15	12	800
5	Lys	Glu	Phe	Val	Phe	*Nph	**Arg**	Leu	10	11	1100	–	–	n.d.	10	1.6	160	6	10	1670

Assays were performed at 37°C in 0.1 M sodium formate buffer pH 4.4 with a final ionic strength of 0.1 M. Nph, *p*-nitrophenylalanine; *, scissile peptide bond; n.d., not determined.

previously [7,8] and values for k_{cat} were calculated from the equation $V_{max} = k_{cat}[E]$.

K_i values were determined at pH 4.4 as described previously [7,8] using the chromogenic substrate Leu-Glu-Arg-Ile-Phe*Nph-Ser-Phe for the plasmepsins [5] and Lys-Pro-Ile-Glu-Phe*Nph-Arg-Leu for the human aspartic proteinases [7].

3. Results and discussion

The recombinant and naturally-occurring forms of plasmepsins I and II were assayed for their ability to hydrolyse five synthetic peptide substrates (Table 1). Peptide 1 was based on the sequence surrounding the Phe-33-Leu-34 peptide bond in α-globin which is cleaved to initiate haemoglobin degradation [6]; and peptides 2–5 form part of a rational series of substrates with single amino acid replacements. For both forms of plasmepsin II, similar values were derived for each of the kinetic parameters k_{cat}, K_m and k_{cat}/K_m with substrates 1 and 2 (Table 1). Thus, recombinant and naturally-occurring forms of plasmepsin II appeared to be exerting comparable activities. For this reason, peptides 3–5 were not examined with the naturally-occurring plasmepsin II since supplies of this form of this enzyme were severely limited.

In contrast, the values measured for the specificity constants (k_{cat}/K_m) for the hydrolysis of all five peptide substrates (Table 1) by recombinant plasmepsin I ranged from 3–10-fold lower than those determined for the naturally-occurring form of the enzyme. Closely similar K_m values were detected with all substrates tested, however, so that the differences in k_{cat}/K_m appear to originate predominantly in a disparity in the respective k_{cat} values rather than in substrate binding affinities.

This interpretation was further substantiated when the interaction of two peptide inhibitors was examined with the

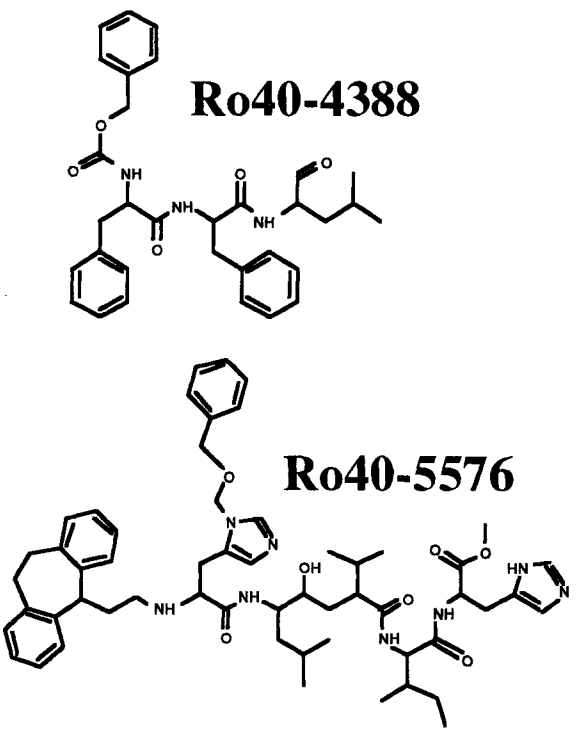

Fig. 1. Inhibitory compounds.

Table 2
Inhibition constants for the interaction of naturally-occurring and recombinant plasmepsins I and II with isovaleryl pepstatin and Ro40-4388

Inhibitor	Plasmepsin II		Plasmepsin I	
	Recombinant, K_i (nM)	Naturally occurring, K_i (nM)	Recombinant, K_i (nM)	Naturally occurring, K_i (nM)
Iva-Val-Val-Sta-Ala-Sta	<0.1	<0.1	0.7	0.5
Ro40-4388	500	500	4	9

Assays were performed at 37°C in 0.1 M sodium formate buffer pH 4.4 with a final ionic strength of 0.1 M. Precise values for isovaleryl pepstatin with plasmepsin II could not be determined due to problems with mutual depletion relative to the concentration of plasmepsin II that needed to be included in the enzyme assay that was used.

recombinant and naturally-occurring forms of plasmepsin I. Closely similar values for K_i were measured for both forms of plasmepsin I when the interaction was measured with either (i) a tight binding inhibitor (pepstatin), or (ii) a less potent inhibitor (Ro40-4388) (Fig. 1; Table 2). Thus it would seem that binding of substrates and inhibitors (as reflected in K_m and K_i respectively) is comparable in both recombinant and naturally-occurring forms of plasmepsin I and therefore it would appear to be justifiable to use the recombinant form of plasmepsin I in studies to enable the design and development of novel inhibitors as anti-malarial drugs. However, the efficiency of substrate cleavage (as reflected in k_{cat} and k_{cat}/K_m) does appear to be compromised in the recombinant form of plasmepsin I relative to the naturally-occurring form. This effect is unlikely to be due to differential glycosylation of recombinant and naturally-occurring forms since recent studies have shown that the plasmepsins are not glycosylated [9].

A somewhat similar situation has been described previously [10] for the human aspartic proteinase precursor, procathepsin D. Upon exposure to acid in vitro, this zymogen undergoes self-processing to generate an activation intermediate (called pseudocathepsin D) which is incapable of further autocleavage to generate the mature form of cathepsin D (an extrinsic enzyme activity is required for this purpose). Nevertheless, pseudocathepsin D is active towards peptide substrates but it has been shown to have k_{cat} values that are lowered relative to those of mature cathepsin D. It retains 18 of the propart residues that should normally be removed during the activation process and thus consists of mature cathepsin D extended at its N-terminus by a further 18 residues (Fig. 2).

Edman degradation of the recombinant form of plasmepsin I through 16 cycles generated the unequivocal sequence: Phe-Lys-Thr-Gly-Leu-Thr-Gln-Lys-Pro-His-Leu-Gly-**Asn-Ala-Gly-Asp**~.

The parts of the sequence shown in bold represent the N-terminal residues that were reported previously for the naturally-occurring form of the mature enzyme that was isolated from the digestive vacuoles of *P. falciparum* [2]. Thus, the recombinant plasmepsin I that was generated by autoactivation under acid conditions (pH 4.4) in vitro of the (mutant) semi-proplasmepsin I precursor appears to be longer than its natural counterpart by 12 residues.

How would the 12 remaining residues of the propart of plasmepsin I affect its activity relative to naturally-occurring plasmepsin I? In the crystal structures of porcine pepsinogen [11] and human progastricsin [12], a conserved Lys residue at position 36P in the propart appears to interact electrostatically via its positively charged side chain with the carboxylate groups of the two catalytic Asp residues. In pseudocathepsin D, Lys-36P is one of the residues which remains attached to the partially-activated form of the enzyme. It is located 11 residues upstream of the authentic N-terminus and continued interaction of this residue with the catalytic apparatus has been postulated to contribute to the lower k_{cat} values measured for pseudocathepsin D [10]. Alignment of the extended N-termini of pseudocathepsin D and the activated recombinant plasmepsins generated in this study (Fig. 2) shows that, like pseudocathepsin D, recombinant plasmepsin I has a Lys residue 11 amino acids upstream of the authentic N-terminus of the naturally-occurring protein. This may well be in a position to be involved in interactions at the active site resulting in the decrease that was observed in the k_{cat} values measured for the recombinant form of plasmepsin I.

For comparative purposes, the N-terminal sequence of recombinant plasmepsin II (produced correspondingly by autoactivation of the semi-proplasmepsin II precursor at pH 4.4) was also determined and two overlapping sequences were detected. The major sequence (62%) was found to be Leu-Gly-**Ser-Ser-Asn-Asp-Asn-Ile-Glu-Leu-Val-Asp-Phe-Gln-Asn**~ whilst the second sequence (38%) was Leu-Asn-Ser-Gly-Leu-Thr-Lys-Thr-Asn-Tyr-Leu-Gly-**Ser-Ser-Asn-Asp**~. In both cases, the bold sequence represents the N-terminus of mature plasmepsin II as reported for the enzyme isolated from *P. falciparum* [2]. Thus, following autoactivation of the recombinant semi-proplasmepsin II precursor at pH 4.4 in vitro, the predominant form of recombinant plasmepsin II produced was closely similar to its naturally-occurring counterpart, being longer at its N-terminus by just two residues. However,

```
PM I:                                  Phe-Lys-Thr-Gly-Leu-Thr-Gln-Lys-Pro-His-Leu-Gly-Asn-Ala~

PM II:                                 Leu-Asn-Ser-Gly-Leu-Thr-Lys-Thr-Asn-Tyr-Leu-Gly-Ser-Ser~

Ps CatD:   Ile-Ala-Lys-Gly-Pro-Val-Ser-Lys-Tyr-Ser-Gln-Ala-Val-Pro-Ala-Val-Thr-Glu-Gly-Pro~
```

Fig. 2. Alignment of the extended N-termini of autoactivated recombinant forms of plasmepsin I and plasmepsin II with that of human pseudocathepsin D. N-terminal sequences are shown for the autoactivated form of recombinant plasmepsin I (PM I), the longer activation product of recombinant plasmepsin II (PM II) and pseudocathepsin D (Ps CatD). Residues corresponding to the N-termini of the naturally-occurring forms of each enzyme are shown in bold. Lys-36P of pseudocathepsin D and equivalent residues in the plasmepsins are underlined.

Table 3
Interaction of four inhibitor compounds with the recombinant plasmepsins and the human aspartic proteinases cathepsin D, cathepsin E, pepsin and gastricsin

Inhibitor	Plasmepsin I, K_i (nM)	Plasmepsin II, K_i (nM)	Cathepsin D, K_i (nM)	Cathepsin E, K_i (nM)	Pepsin, K_i (nM)	Gastricsin, K_i (nM)
Ro40-4388	4	500	70	400	8000	>20 000
Ro40-5576	6	250	2000	5000	3000	5000

Measurements were performed at 37°C with the ionic strength maintained at 0.1 M. Individual assay conditions were as follows: (i) Recombinant plasmepsins I and II; 0.1 M sodium formate buffer pH 4.4. (ii) Recombinant human cathepsin E and naturally-occurring human cathepsin D, pepsin and gastricsin; 0.1 M sodium formate buffer pH 3.1.

a minor (longer) form, similar in length to the recombinant form of plasmepsin I, was also found to be present in the mixture. In the 12 residue extension of this longer form of plasmepsin II, a lysine residue is not found at the corresponding position to that of recombinant plasmepsin I and pseudocathepsin D (Fig. 2). Instead, an uncharged Asn residue replaces the Lys-36P, thus precluding any detrimental interactions with the residues of the catalytic apparatus. Thus, it is perhaps not surprising that the kinetic parameters measured for substrate hydrolysis (Table 1) and the binding of the two inhibitors, isovaleryl pepstatin and Ro40-4388 (Table 2) are closely similar for both recombinant and naturally-occurring forms of plasmepsin II.

Since the activity of recombinant plasmepsin II can be taken as representative of its naturally-occurring counterpart, the data in Table 1 also allow comparison of this plasmepsin II with naturally-occurring plasmepsin I. The haemoglobin-based substrate (peptide 1) is hydrolysed readily by naturally-occurring plasmepsin I and both forms of plasmepsin II. When kinetic constants are compared between recombinant plasmepsin II and naturally-occurring plasmepsin I for a series of peptides [2–5] small differences in specificity emerge. When the aromatic Phe in the P3 position of peptide 2 is replaced by a Leu, the k_{cat} for plasmepsin I decreases 3-fold whereas the k_{cat} for plasmepsin II decreases by less than 2-fold with the result that in terms of the specificity constant (k_{cat}/K_m), recombinant plasmepsin II is more than twice as efficient at hydrolysing peptide 3 than naturally-occurring plasmepsin I.

When the Val in the P2 position in peptide 2 is replaced by Asn (peptide 4), there is a 2-fold decrease in the value of k_{cat}/K_m obtained with naturally-occurring plasmepsin I. This is in contrast to an 8-fold drop apparent for plasmepsin II. Thus, in terms of k_{cat}/K_m values, naturally-occurring plasmepsin I is three times more effective at hydrolysing this substrate with a P2 Asn substituent than plasmepsin II. These differences in substrate specificity can be rationalised by reference to the molecular model of plasmepsin I [5] and the crystal structure of plasmepsin II [13] which indicate that plasmepsin I has larger S3 and S2 subsites, and is better able to accommodate P3 Phe and P2 Asn residues than plasmepsin II.

This distinction in subsite requirements for the hydrolysis of substrates was also reflected in the susceptibility of both enzymes to inhibition. The general aspartic proteinase inhibitor isovaleryl pepstatin is highly potent against naturally-occurring plasmepsin II and less active against naturally-occurring plasmepsin I (Table 2). The converse is true for the inhibitor Ro40-4388 which has at least a 50-fold higher affinity for plasmepsin I than for plasmepsin II. For each inhibitor the K_i value measured for recombinant plasmepsin II was indistinguishable from that calculated for its naturally-occurring counterpart (Table 2). In addition, despite the complexities described above for specificity constants of naturally-occurring and recombinant plasmepsin I, comparable values were measured for both forms of this enzyme.

To consider compounds such as Ro40-4388 as leads for drug developments, selectivity for the target enzyme is essential. Thus, the selectivity of Ro40-4388 was assessed by examining its ability to inhibit the well-documented human aspartic proteinases cathepsin D, cathepsin E, pepsin and gastricsin (Table 3). Ro40-4388 is a good inhibitor of plasmepsin I (Table 2; [5]) but it also has a relatively poor selectivity since the K_i that was measured against cathepsin D was only approximately 17-fold poorer than that determined for plasmepsin I (and its affinity for both cathepsin D and cathepsin E was greater than for plasmepsin II (Table 3)). Therefore, a second compound Ro40-5567, which inhibited plasmepsin I with efficiency comparable to Ro40-4388, was tested with the human enzymes (Table 3). This inhibitor shows at least a 300-fold greater affinity for plasmepsin I than for any of the human enzymes. Although this would be considered insufficient selectivity for a drug, it nonetheless represents a considerable improvement over Ro40-4388. Ro40-5567 has also been shown to kill parasites in culture [5], thus an understanding of its interaction with plasmepsin I may be useful in the future development of plasmepsin inhibitors as drugs.

Acknowledgements: This investigation received financial support from the UK Biotechnology and Biological Sciences Research Council (for L.T.), from the Royal Society (for C.B.), from the United Nations Development Programme/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR) (for J.K. and C.B.) and from the National Institutes of Health (AI-37977 to D.E.G.; AI-39211 to J.W.). D.E.G. is a recipient of the Burroughs Wellcome Fund Scholar Award in Molecular Parasitology.

References

[1] 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) EMBO J. 13, 306–317.
[2] Gluzman, I.Y., Francis, S.E., Oksman, A., Smith, C.E., Duffin, K.L. and Goldberg, D.E. (1994) J. Clin. Invest. 93, 1602–1608.
[3] Rosenthal, P.J., McKerrow, J.H., Aikawa, M., Nagasawa, H. and Leech, J.H. (1988) J. Clin. Invest. 82, 1560–1566.
[4] Hill, J., Tyas, L., Phylip, L.H., Kay, J., Dunn, B.M. and Berry, C. (1994) FEBS Lett. 352, 155–158.
[5] Moon, R.P., Tyas, L., Certa, U., Rupp, K., Bur, D., Jaquet, C., Matile, H., Loetscher, H.-R., Grueninger, F., Kay, J., Dunn, B.M., Berry, C. and Ridley, R.G. (1997) Eur. J. Biochem. 244, 552–560.
[6] Goldberg, D.E., Slater, A.F.G., Beavis, R., Chait, B., Cerami, A. and Henderson, G.B. (1991) J. Exp. Med. 173, 961–969.
[7] Jupp, R., Dunn, B.M., Jacobs, J.W., Vlasuk, G., Arcuri, K.E., Verber, D.F., Perlow, D.S., Payne, L.S., Boger, J., de Laszlo, S.,

Chakravarty, P.K., tenBroeke, J., Hangauer, D.G., Ondeyka, D., Greenlee, W.J. and Kay, J. (1990) Biochem. J. 265, 871–878.
[8] Hill, J., Montgomery, D.S. and Kay, J. (1993) FEBS Lett. 326, 101–105.
[9] Francis, S.E., Banerjee, R. and Goldberg, D.E. (1997) J. Biol. Chem. 272, 14961–14968.
[10] Beyer, B.M. and Dunn, B.M. (1996) J. Biol. Chem. 271, 15590–15596.
[11] James, M.N.G. and Sielecki, A.R. (1986) Nature 319, 33–38.
[12] Moore, S.A., Sielecki, A.R., Chernaia, M.M., Tarasova, N.I. and James, M.N.G. (1995) J. Mol. Biol. 247, 466–485.
[13] Silva, A.M., Lee, A.Y., Gulnik, S.V., Majer, P., Collins, J., Bhat, T.N., Cachau, R.E., Luker, K.E., Gluzman, I.Y., Francis, S.E., Oksman, A., Goldberg, D.E. and Erickson, J.W. (1996) Proc. Natl. Acad. Sci. USA 93, 10034–10039.