



Bioinformatics based discovery of new keratinases in protease family M36

Qiu, Jingwen; Barrett, Kristian; Wilkens, Casper; Meyer, Anne S.

Published in:
New Biotechnology

Link to article, DOI:
[10.1016/j.nbt.2022.01.004](https://doi.org/10.1016/j.nbt.2022.01.004)

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Qiu, J., Barrett, K., Wilkens, C., & Meyer, A. S. (2022). Bioinformatics based discovery of new keratinases in protease family M36. *New Biotechnology*, 68, 19-27. <https://doi.org/10.1016/j.nbt.2022.01.004>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

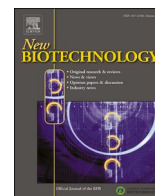
- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Contents lists available at ScienceDirect

New BIOTECHNOLOGY

journal homepage: www.elsevier.com/locate/nbt

Full length Article

Bioinformatics based discovery of new keratinases in protease family M36

Jingwen Qiu, Kristian Barrett, Casper Wilkens, Anne S. Meyer *

Section for Protein Chemistry and Enzyme Technology, Department of Biotechnology and Biomedicine, DTU Bioengineering, Building 221, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Keywords:

M36 protease family
Keratin
Conserved peptides
Proteases heterologous expression
Keratinase kinetics
Keratinase thermal stability

ABSTRACT

Keratinases are proteases that can catalyze the degradation of insoluble keratinous biomass. Keratinases in protease family M36 (MEROPS database) are endo-acting proteases. In total, 687 proteases are classified in family M36. In the present study, new keratinolytic enzymes were identified in protease family M36 using the bioinformatics tool Conserved Unique Peptide Patterns (CUPP). Via CUPP, M36 family members were classified into 11 groups, with CUPP group 1 containing the three currently known and sequenced family M36 keratinases (derived from the fungi *Fusarium oxysporum*, *Microsporium canis* and *Onygena corvina*) as well as an additional 71 uncharacterized M36 proteases. In order to assess the relevance of CUPP group 1 categorization to keratinolytic function, four uncharacterized M36 proteases and the known keratinase from *F. oxysporum* (in CUPP group 1) were selected for recombinant expression and keratinolytic activity assessment. The four hitherto unknown M36 proteases were from *Phaeosphaeria nodorum*, *Aspergillus clavatus*, *Pseudogymnoascus pannorum* and *Nectria haematococca*, and represent four different fungal taxonomical classes. The genes encoding the selected M36 proteases were individually expressed in *Pichia pastoris* and all proteases displayed keratinase activity on keratin azure. Additionally, the activity on different keratinase substrates, optimal reaction conditions and thermal stability were determined for the two most active new keratinases. The results validate the applicability of CUPP for function-based discovery of non-characterized keratinases and present new robust keratinases for potential use in keratin upgrading.

Introduction

Keratin is an abundant, insoluble, fibrous protein that constitutes the structural protein of mammalian horns, wool and claws (α -keratin), and feathers, beaks, and reptile shells (β -keratin). Keratin forms highly stable materials due to its tight packing, disulfide bonds and hydrophobic interactions, and resists degradation by conventional proteolytic enzymes such as pepsin and trypsin [1]. However, keratinases (EC 3.4.-.-) can catalyze hydrolytic degradation of keratin to produce amino acids and/or small soluble peptides [2]. Recently, it has been recognized that microbial keratinases may be used for utilization of keratin via controlled enzymatic degradation to create peptides and amino acid products for applications in functional skin care products, animal feed, or fertilizers [3,4]. Keratin is an abundant material, e.g. the annual global chicken feather waste from poultry processing was estimated to

amount to more than 4.7 million tons per year in 2019 [5]. Likewise, horns, beaks, and pig bristles are keratin-rich side products of meat production [6]. Discovery of robust microbial keratinases is an important first step for development of new bioprocesses for keratin utilization.

Important keratinase-secreting bacteria include e.g. *Bacillus* sp. [7] and *Streptomyces* sp. [8,9], while keratinase-secreting fungi include dermatophytic species such as *Trichophyton rubrum* and *Microsporium canis*, as well as non-pathogenic fungi such as *Onygena corvina* [10,11]. Keratinases are no longer categorized with a full 4-digit EC number, and several keratinase sequences are available in the NCBI database. The known sequenced keratinases belong to at least 14 different protease families, namely M3, M4, M14, M16, M28, M32, M36, M38, M55, S1, S8, S9, S10, and S16 [11] (MEROPS database [12]). They may act via endo-attack (e.g. proteases belonging to families M4, M16, M36, S1, S8,

Abbreviations: AcMep, M36 protease derived from *Aspergillus clavatus*; AfuMep, M36 protease from *Aspergillus fumigatus*; CUPP, Conserved Unique Peptide Patterns; FoMep, M36 protease derived from *Fusarium oxysporum*; McMep, M36 protease derived from *Microsporium canis*; NhMep, M36 protease derived from *Nectria haematococca*; OcMep, M36 protease derived from *Onygena corvina*; PnMep, M36 protease derived from *Phaeosphaeria nodorum*; PpMep, M36 protease derived from *Pseudogymnoascus pannorum*.

* Corresponding author.

E-mail address: asme@dtu.dk (A.S. Meyer).

<https://doi.org/10.1016/j.nbt.2022.01.004>

Received 5 August 2021; Received in revised form 8 January 2022; Accepted 11 January 2022

Available online 12 January 2022

1871-6784/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

S16), exo-attack (proteases in families M14, M28, M38, M55, S9, S10) or on oligopeptides only (the proteases in families M3, M32) [11,13–15].

The M36 family proteases have been confirmed as endo-acting and have been shown to be able to catalyze cleavage of extracellular matrix proteins, such as elastin and keratin [12]. Currently, those from the fungi *Fusarium oxysporum* [16], *M. canis* [17] and *O. corvina* [10] are known to catalyze degradation of keratin. However, compared to the large number of keratinases described in family S8 [11], only a limited number of M36 proteases have been shown to have keratinolytic activity and there is very little information on the specificity, optimal reaction conditions, and kinetics of these enzymes. The aim here was to discover new M36 keratinases, express them recombinantly and validate their enzymatic activity on keratinous substrates.

Through a phylogenetic approach, homologous amino acid sequences can be analyzed to allow reliable extrapolation between proteins of known function to those of unknown function [18,19]. However, for very large protease families the construction of phylogenetic trees using multiple sequence alignment requires a significant amount of computer power [20]. In contrast, the new bioinformatics based enzyme discovery tool Conserved Unique Peptide Patterns (CUPP), has the ability to handle large data sets efficiently [21,22]. CUPP is a peptide-based similarity assessment algorithm that can group proteins according to peptide motif resemblance and thus facilitate detailed sub-grouping of enzymes within protein families or subfamilies [22]. Recently, CUPP has successfully been applied for classification and annotation of carbohydrate-active enzymes (CAZy) [21,22], but until now, it has not been explored for classification of proteases. An important part of our research objective was to assess the use of the CUPP tool to classify proteases and help disclose new M36 keratinases.

Materials and methods

Keratinases and M36 proteases sequence acquisition

The protein sequences of the three currently known M36 keratinases FoMep (NCBI accession no. BAM84176), McMep (NCBI accession no. CAD35288) and OcMep (NCBI accession no. AJD23141) derived from *F. oxysporum* [16], *M. canis* [17] and *O. corvina* [10], respectively, were downloaded from the NCBI database. All 687 proteases in the M36 protease family were obtained from the MEROPS database (February 22nd, 2019). The bioinformatics analysis was conducted as described in Suppl. Fig. S1. The M36 members and the three known keratinase sequences were clustered at 90 % sequence threshold using CD-HIT [23]. CD-HIT is an incremental algorithm that is used to remove redundancy; it also reduces storage space, computational time and noise interference in further bioinformatics analyses [23].

Sequence alignment and phylogenetic tree construction

For phylogenetic tree construction, the 508 representative protease sequences from the CD-HIT result were aligned in CLC main workbench (version 8.0) by progressive alignment. The protease sequences that were used for phylogenetic analysis and CUPP categorization were without signal peptides and propeptides. The aligned M36 proteases and keratinases were submitted to CIPRES using the RAxML black-box model with substitution matrix LG for phylogenetic analysis [24]. The result was uploaded to iTOL for visualization [25].

CUPP clustering

The representative M36 members remaining after CD-HIT pruning were then used for CUPP clustering using default settings [22], principally as follows. (1) All 508 sequences were subjected to the CUPP analysis [21]. (2) Identification of peptides: CUPP was set to run five iterations of incremental clustering using increasingly more conserved peptides, leading to a conservation fraction of 0.4 [22]. (3)

Identification of CUPP groups: at least five members, each with more than 30 positions covered, could become a member of a CUPP group. (3) Peptides present in at least 20 % of the members of a CUPP group were considered conserved. (4) Finally, an all proteins dendrogram and a protein groups dendrogram was exported from CUPP and visualized in iTOL [25].

Homology modeling of the Pseudogymnoascus pannorum (PpMep) and the Aspergillus clavatus (AcMep) M36 proteases

The catalytic domain structural models of the selected M36 proteases PpMep and AcMep (MEROPS identifiers MER0858341 and MER0086396, respectively) were prepared using HHpred [26] and based on the model of the crystal structure of AfuMep from *Aspergillus fumigatus* (PDB ID 4K90) [27]. The PpMep and AcMep homology models were judged by ProQ [28] and prediction server Zlab [29]. The structures were visualized via PyMol 2.3.3 (Schrödinger, LLC, New York, NY, USA).

Cloning and expression of proteases

The signal peptides of the four putative keratinases AcMep, PpMep, PnMep, NhMep (MEROPS identifiers: MER0086396, MER0858341, MER0081504, MER0295893) and of the benchmark enzyme FoMep (NCBI accession no. BAM84176) were predicted using SignalP 5.0 [30]. Their codon optimized sequences (including an N-terminal 6×HisTag and the propeptide sequences of each enzyme, but excluding the predicted signal peptide sequences) were synthesized and cloned into the pPICZα-A plasmid (GenScript, Piscataway, NJ, USA). The plasmids were propagated in *Escherichia coli* DH5α under Zeocin selection (Invitrogen, Carlsbad, CA, USA), then purified and linearized using MssI (New England BioLabs, Ipswich, MA, USA) [31] and transformed into *Pichia pastoris* X-33 (Invitrogen). For each target protease a colony was selected after 3 d incubation at 30 °C on yeast peptone dextrose plates supplemented with 100 µg mL⁻¹ Zeocin. *P. pastoris* transformants were grown in shake flasks in buffered glycerol-complex medium (BMGY; 28 °C, pH 6, 20 h), and harvested (8000 g, 10 min, 20 °C). The cells were then re-suspended in a buffered methanol-complex medium (BMMY) to a final OD₆₀₀ = 1, and incubated at 20 °C for 72 h with methanol supplementation to 0.5 % (v/v) every 24 h. Then, the cells were harvested (8000 g, 10 min, 20 °C), supernatants were filtrated by Vivaspin filter centrifugation (Vivaspin 20; 30,000 MWCO, GE Healthcare) at 3500 g and the supernatants were buffer changed using 10 mM Tris-HCl buffer pH 7.5 and centrifuged repetitively at 3500 g until the filtrate was colorless. All steps were carried out at 4 °C, and the supernatants were stored at -20 °C.

The predicted molecular weight of the target enzyme was calculated via ProtParam [32]. The purity of the recombinantly produced proteins was determined on SDS-PAGE gels (4–20 %). The protein concentration was determined by the bicinchoninic acid (BCA) method using the Pierce® BCA Protein Assay Kit (Thermo Scientific, Rockford, USA) using bovine serum albumin as standard.

Keratin hydrolysis test

To evaluate the keratinolytic ability of each of the putative keratinases (FoMep, AcMep, PpMep, PnMep, and NhMep), the amount of azure dye released by each protease was assessed using 1 % w/v keratin azure as substrate (Sigma Aldrich, Merck) in extended assay reactions (1 h) at 30 °C, pH 7.5 (50 mM Tris-HCl buffer) in a total reaction volume of 250 µL; the enzymes were added at concentrations 0.3–1 mg/mL due to their different activities. Following the same procedure, pepsin and trypsin were also examined for their ability to hydrolyze keratin azure using a ratio of enzyme:substrate = 1:2000 (w/w) [33] at 30 °C, for trypsin at pH 7.5 (Tris-HCl buffer) and for pepsin at pH 2 (Glycine-HCl buffer). For the assay with addition of dithiothreitol (DTT), the keratin

azure was pretreated with 2 mM DTT for 2 h at 30 °C, then the DTT concentration was decreased to 0.4 mM for the enzymatic hydrolysis reaction by dilution via addition of buffer and protease. The control was treated in the same way as the sample, but using heat-inactivated (95 °C, 5 min) proteases. For these measurements, the supernatant samples were centrifuged immediately (12,000 g; 3 min; 4 °C) after 1 h reaction, and the absorbance of the total released azure was measured at 595 nm (A_{595}). All reactions were run in triplicate, and the data are reported as total A_{595} /mg protein.

Optimal reaction conditions study by response surface modeling

A 3-level full factorial design was performed to identify the pH-temperature reaction optimum and study the influence of pH and temperature on the rate of keratinase catalyzed keratin azure hydrolysis. Based on preliminary experiments (data not shown), the factor levels were set to pH 8–10 (50 mM Tris-HCl buffer-50 mM CAPS buffer) and 25–45 °C, with pH 9, 35 °C as repeated center point to give a total of nine different reaction combinations in 12 experiments for each keratinase (all in triplicate).

In all reactions 1.0 % (w/v) keratin azure substrate was suspended in the buffer (pH according to the design) and pre-incubated at the assigned temperature for 3 min before adding the keratinase (tested at dosage levels of 0.007–0.015 mM due to their different activities). For each reaction, the initial rate was determined from linear regression of the initial data points (within 10 min of reaction). For each of the pH-temperature conditions for each enzyme, measurements were conducted by sampling the supernatant for absorbance measurement at the set time point; sampled supernatants were centrifuged immediately (12,000 g, 4 °C), and the A_{595} was measured on 200 μ L of the clarified supernatant, in a Multiskan GO microplate spectrophotometer (Thermo Fisher Scientific, USA). The data were calculated into activity units via the Beer–Lambert law assuming a molar absorption coefficient of the azure of 135,000 L/(mol·cm) [34]. One keratinase unit was defined as μ M azure product released per minute at the designed reaction conditions compared with the control reaction. Heat inactivated enzyme (treated at 95 °C for 5 min) was used as control.

The statistical design program MODDE 12.01 (Umetri AB, Umeå, Sweden) was used as an aid to design the factorial experiments and to fit and analyze the data by multiple linear regression. Significance of the results was established at $p \leq 0.05$.

Kinetics data

The kinetic parameters were estimated for FoMep, PpMep, and AcMep on keratin azure. At least seven different keratin azure substrate concentrations from 0.2 % to 16.0 % (w/v) were used, in assay reactions as described above at the optimum reaction conditions defined from the RSM data for each keratinase; FoMep pH 8.6, 38 °C; PpMep pH 8.4, 40 °C; AcMep pH 8.3, 45 °C. The kinetic parameters were calculated using the Michaelis–Menten equation in Origin (version OriginPro 2018b).

Thermal stability of the new fungal keratinases

The keratinases were incubated at different temperatures of 30, 40, 50 and 60 °C, for defined time periods (1–60 min. in 50 mM Tris-HCl buffer). The residual activity was determined by the keratin azure assay at optimum reaction conditions as described above. The first order inactivation rate constant (k_D) was obtained from: $\ln(\text{activity}) = \ln(\text{activity}_0) - k_D t$, where t designates the incubation time of the enzyme at the particular temperature. The half-life was calculated from $t_{1/2} = \ln 2 / k_D$, i.e. defined as the time when the residual activity of the enzyme is reduced to half of its original activity after incubation at a certain temperature.

Activity assessment on different substrates

The activity of each enzyme was compared on soluble azocasein (Megazyme, Bray, Ireland), insoluble keratin azure and azokeratin, respectively, (the azokeratin prepared from pig bristles and hooves [10, 35,36], was kindly provided by Professor Søren Sørensen, University of Copenhagen, Denmark). For measuring the proteolytic activity on azocasein, 1 % w/v azocasein was suspended in 50 mM Tris-HCl buffer and each enzyme, FoMep, PpMep, and AcMep, was added (dosage levels ranging from 0.007–0.015 mM). The reactions were stopped after 10 min by adding trichloroacetic acid (TCA) to a final concentration of 0.2 M, and then incubated at 4 °C for 30 min for protein precipitation. Each mixture was then centrifuged (12,000 g; 5 min; 4 °C) to remove the substrate. 100 μ L supernatant was then immediately transferred to a microliter plate containing 25 μ L NaOH (final concentration 0.36 M NaOH). For the control runs, TCA was added before adding the enzyme. On azokeratin, the assays were run as described for the keratin azure assay, except that azokeratin was used as substrate (1 % w/v concentration in the final assay). All reactions were run in triplicate at the optimal conditions for each enzyme (FoMep pH 8.6, 38 °C; PpMep pH 8.4, 40 °C; AcMep pH 8.3, 45 °C).

The activity was quantified from spectrophotometric measurements by calculating the molar concentration of azo dye released according to the Beer–Lambert law, assuming a molar extinction coefficient of 82,600 L/(mol·cm) at 420 nm for the azocasein assay (that includes addition of NaOH prior to absorbance measurement) [37] and 19,600 L/(mol·cm) at 415 nm for the azokeratin assay [35,38]. One enzyme unit was defined as μ M azo dye released per min under the test conditions compared with the control reaction.

Results and discussion

Phylogenetic analysis of family M36 proteases

The 508 proteases (from CD-HIT at 90 % sequence identity threshold), including the three known keratinolytic proteases FoMep, McMep, and OcMep in the M36 family, were subjected to alignment through CLC main workbench, and used to construct a maximum likelihood phylogenetic tree (Fig. 1a and Suppl. Fig. S2). Because of the diversity of the 508 sequences, it is difficult to identify the large clades supported by high bootstrap values directly in the phylogenetic tree. Moreover, based only on the tree, it would be difficult to systematically predict new putative keratinases. Therefore, CUPP clustering was used.

CUPP clustering of family M36 proteases

CUPP was applied to cluster the 508 representative sequences in protease family M36 resulting after CD-HIT reduction. A total of 455 proteins of the 508 representative sequences were assigned to a CUPP group (Table 1, Suppl. Table S1). The 53 extra proteins were not assigned, either being singletons (and hence not constituting a group) or because they had too few positions covered to be included in a CUPP group [22]. The CUPP analysis resulted in 11 CUPP groups of conserved peptides (Table 1). CUPP group 1 included the three known keratinolytic enzymes FoMep, McMep, OcMep in addition to another 71 proteases that have not been described as having keratinolytic abilities (Table 1, Suppl. Table S1). The group 1 proteases are derived from fungi belonging to four different taxonomical classes, namely *Eurotiomycetes*, *Sordariomycetes*, *Dothideomycetes* and *Leotiomycetes* (Suppl. Table S1).

Systematic bioinformatics based keratinase prediction

The CUPP results were entered into the phylogenetic tree, and the 11 different CUPP groups of M36 proteases are shown in Fig. 1a. Members of the same group were generally found to have short intervening distances in the tree (Fig. 1a, Suppl. Fig. S2). For instance, CUPP group 1

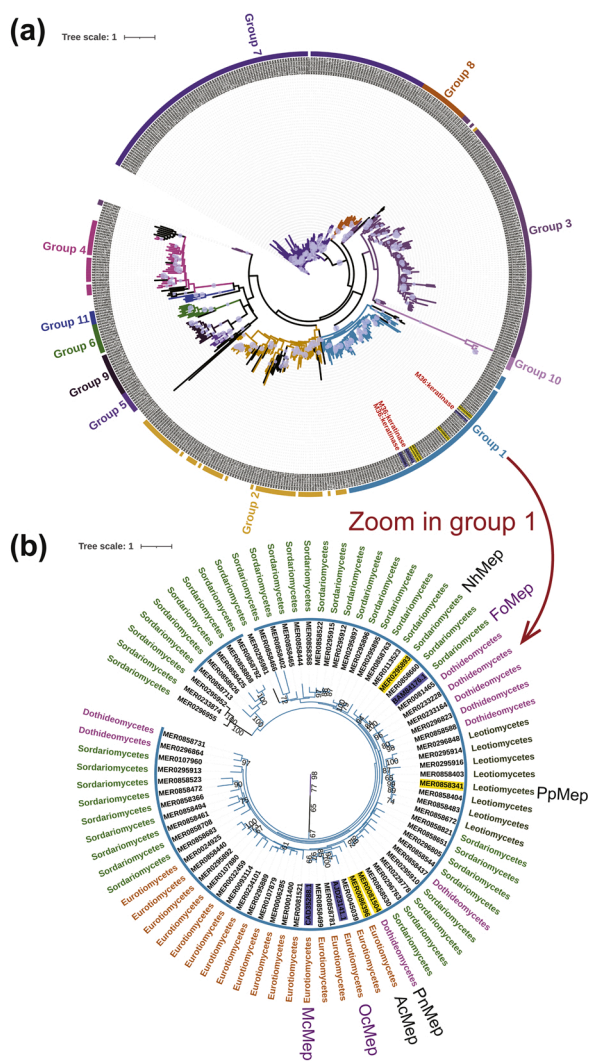


Fig. 1. Phylogenetic trees of M36 proteases. (a) Phylogenetic tree of the representative M36 proteases and three M36 keratinases: FoMep derived from *Fusarium oxysporum*, McMep from *Microsporium canis* and OcMep from *Onygena corvina*. Tree construction was by the maximum likelihood method. Inside the tree are 11 numbered rings that correspond to 11 CUPP groups. The three M36 keratinases (FoMep: NCBI accession number BAM84176, McMep: CAD35288 and OcMep: AJD23141) are labeled as “M36: keratinases”. The dots in the tree correspond to branches with bootstrap support $\geq 60\%$. The bootstrap value $\geq 60\%$ is shown in numbers in Suppl. Fig. S2. (b) Pruned phylogenetic tree for members in M36 protease family CUPP group 1. All the original organisms in different taxonomical classes are displayed in various colors. The accession numbers with a purple background in the tree are three confirmed M36 keratinases, while yellow highlighted members (MER0295893, MER0858341, MER0081504, MER0086396 corresponding to NhMep from *N. haematococca*, PpMep from *P. pannorum*, PnMep from *P. nodorum* and AcMep from *A. clavatus*, respectively) were selected for experimental characterization of keratinase activity. Bootstrap is shown as numbers in the tree.

corresponded with the tree cluster with bootstrap support value of 67 (Fig. 1b). Similarly, other CUPP groups were also found to match phylogenetic clusters that had bootstrap values >50 . Additionally, the CUPP grouping was able to identify new and smaller clusters, and the results thus provided a higher resolution to distinguish diversity. For example, CUPP groups 5, 10 and 11 contained only 8, 5 and 5 members, respectively (Table 1, Fig. 1a).

It is noteworthy that not all the members of the clade, such as the

Table 1
Summary of M36 protease family CUPP groups.

CUPP group number	CUPP group members	Numbers of keratinases	Taxonomy		
			Taxonomical classes	Bacteria	Fungi
1	74	3	4		×
2	64	0	7	×	
3	96	0	15	×	×
4	26	0	3	×	
5	8	0	4	×	
6	11	0	3	×	
7	130	0	1		×
8	19	0	1		×
9	16	0	1	×	
10	5	0	2	×	
11	5	0	2	×	
Total	455	3	33	×	×

four proteases MER0295955, MER0295952, MER0233874, MER0295981 (Fig. 1b), were classified into a CUPP group even though they were close to group 1 members in the same clade (bootstrap = 67). Similarly, some proteases near CUPP groups 2 and 4 were not grouped by CUPP (Fig. 1a) since their similarity to any group was too low [22]. Group 1 contained all the three known keratinolytic proteases as members, thus was the only group investigated further.

Since CUPP clusters proteins according to similarity of conserved peptide motifs, members of the same group may share a similar molecular function or functional feature [22]. This suggests that the 71 M36 proteases in group 1 (containing the three known keratinases) (Table 1) have the highest possibility of being keratinolytic. In order to discover new M36 keratinases and validate the relevance of CUPP group 1 to keratinolytic function, four candidates of unknown function in M36 CUPP group 1 were selected for keratinase activity testing, namely the four fungal proteins PpMep (*Pseudogymnoascus pannorum*), AcMep (*Aspergillus clavatus*), PnMep (*Phaeosphaeria nodorum*), and NhMep (*Nectria haematococca*) (Table 2, Fig. 1b). To confirm that the function of the keratinase candidates was similar despite their taxonomical diversity, the four candidates individually represented four different taxonomical classes (along with FoMep [16] as the benchmark).

Eight M36 proteases in CUPP group 1 were from the class *Leotiomycetes*; five of these originated from *P. pannorum* and of these MER0858341 (PpMep) was selected. No keratinolytic proteins have been reported from this species. There are likewise no reports that *A. clavatus* [39] (AcMep, MER0086396), class *Eurotiomycetes*, can degrade keratin, although several other members of the genus *Aspergillus* in CUPP group 1, such as *A. niger* and *A. fumigatus*, have been extensively studied as keratinolytic organisms [40,41]. *P. nodorum* [42] (PnMep, MER0081504), class *Dothideomycetes*, is known as a major necrotrophic fungal pathogen of wheat [43], but has not been reported to be keratinolytic. Lastly, a candidate from *N. haematococca* [44], class *Sordariomycetes* (NhMep, MER0295893), was selected because it was placed close to keratinase FoMep BAM84176 with a bootstrap value of 94 (Fig. 1b).

Sequence analysis

From the sequence analysis of the four putative M36 proteases including the FoMep as benchmark, it was evident that they all contained the canonical “HEXXH” motif found in metallopeptidases from clan MA [13]. The motif was a conserved HEYTH in all the sequences (dashed box, amino acid 184–188, Fig. 2). The analysis also included the catalytic domain of AfuMep derived from *A. fumigatus* (MER0001400, PDB ID 4K90) because AfuMep served as the model for homology modeling of the selected proteases. An additional active site glutamic acid residue is located downstream of the motif, and was seen in all the selected proteases (residue 214, Fig. 2). Although the sequence identity

Table 2
Known, sequenced M36 keratinolytic enzymes and the new selected putative M36 keratinases.

Protease name	NCBI accession number	MEROPS identifier	Source organism	Taxonomical class	Mw (kDa)	Protein length	Catalytic domain ^c	Reference
M36 keratinases								
FoMep	BAM84176	–	<i>Fusarium oxysporum</i>	<i>Sordariomycetes</i>	~48 ^a	632	245–632	[16]
McMep	CAD35288	–	<i>Microsporium canis</i>	<i>Eurotiomycetes</i>	43.5	633	247–633	[17]
OcMep	AJD23141	MER0858361	<i>Onygena corvina</i>	<i>Eurotiomycetes</i>	43.2 ^b	634	246–634	[10]
New putative M36 proteases with uncharacterized keratinolytic function								
PpMep	–	MER0858341	<i>Pseudogymnoascus pannorum</i>	<i>Leotiomycetes</i>	~44 ^a	638	248–635	–
AcMep	–	MER0086396	<i>Aspergillus clavatus</i>	<i>Eurotiomycetes</i>	~44 ^a	634	247–633	[39]
PnMep	–	MER0081504	<i>Phaeosphaeria nodorum</i>	<i>Dothideomycetes</i>	41.5 ^b	626	242–625	[42]
NhMep	–	MER0295893	<i>Nectria haematococca</i>	<i>Sordariomycetes</i>	42.0 ^b	637	248–637	[44]

^a Estimated from SDS PAGE gels in this work.
^b The molecular weight was calculated via ProtParam [32].
^c The catalytic domain was confirmed based on the confirmed M36 protease AfuMep (derived from *Aspergillus fumigatus*) catalytic domain sequence (chain A, NCBI accession number 4K90_A).

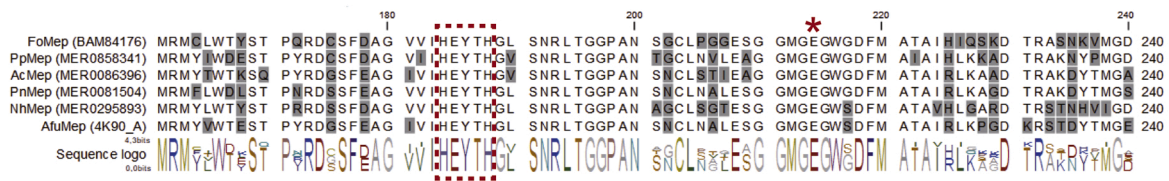


Fig. 2. Partial catalytic domain sequence alignment of selected M36 proteases FoMep (NCBI accession no. BAM84176), PpMep (MER0858341), AcMep (MER0086396), PnMep (MER0081504), NhMep (MER0295893) and the catalytic domain sequence of M36 protease AfuMep (MER0001400; NCBI accession no. 4K90_A) by CLC main workbench 8.0. The metalloprotease motif “HEXXH” is identified within the dashed box. “*” indicates the additional glutamic acid in the active site. The complete catalytic domain sequences alignment is in Suppl. Fig. S3.

was < 90 % (due to the CD-HIT reduction at 90 % sequence identity threshold), a wider comparison of the catalytic domain sequences showed that they had several congruent amino acids (Suppl. Fig. S3). The similarity between each of the selected proteases (catalytic domain) is <75 %, except for NhMep, which has a higher similarity to FoMep (similarity of 86.7 %).

Heterologous expression and activity assessment of new M36 keratinolytic enzymes

All the four selected enzymes, PpMep, AcMep, PnMep, NhMep as well as FoMep were expressed recombinantly in *P. pastoris*. The bands of FoMep and PpMep corresponded to the calculated molecular weights of the mature enzymes, i.e. 48.1 and 43.8 kDa, respectively (Suppl. Fig. S4). The recombinantly produced AcMep was glycosylated, but treatment with EndoH revealed that the molecular weight of approximately 44 kDa was as expected after deglycosylation (Suppl. Fig. S4). The molecular weight data (Table 2) thus indicated that for each enzyme, the pro-peptide was recognized, excised, and degraded during recombinant expression, either by each protease itself or by the protease machinery in *P. pastoris* [45].

All the enzymes exhibited keratinolytic ability on keratin azure, with FoMep, PpMep, and AcMep having higher activity than PnMep and NhMep (Fig. 3). The data also showed that FoMep (2.1 A₅₉₅/mg protein), PpMep (0.4 A₅₉₅/mg protein) and AcMep (1.1 A₅₉₅/mg protein) were able to catalyze the degradation of the keratin azure substrate even without the addition of reducing agent or disulfide reductase (Fig. 3). Higher keratin hydrolysis could be obtained with addition of DTT (Fig. 3), due to reduction of the disulfide bonds providing better access to the proteolytic cleavage sites in the keratin backbone. With AcMep, however, there was no significant effect of adding DTT (Fig. 3). Pepsin

and trypsin were also examined for their ability to catalyze keratin azure degradation; pepsin treatment did not release any detectable products even on DTT treated keratin substrate (data not shown). The keratin hydrolysis of trypsin was 0.25 A₅₉₅/mg protein without addition of DTT after 1 h hydrolysis (Suppl. Fig. S5), which was lower than for FoMep, PpMep and AcMep. With addition of DTT, hydrolysis by trypsin increased to 0.66 A₅₉₅/mg protein (Suppl. Fig. S5). The data agree with a recent report showing keratinolytic activity of trypsin on keratin azure, and that this activity was higher than that of a family S1 keratinase studied (T-like protease) [46]. As the SDS PAGE data revealed (Suppl. Fig. S4), PnMep and NhMep were not highly expressed by *P. pastoris*. Nevertheless, the catalytic keratin hydrolysis ability of PnMep and NhMep was significantly higher than that of a control empty plasmid expression indicating a weak keratinolytic ability of 0.47 A₅₉₅/mg protein of each of these enzymes (Fig. 3).

Overall, the detected keratinolytic ability of the selected M36 proteases indicated that the CUPP grouping was able to identify relevant functional features for keratinase activity based on conserved peptide patterns. The results imply that the CUPP tool can indeed assist in identifying new keratinolytic enzymes. The current study thus validates the applicability of CUPP to predict new functional proteases. FoMep, PpMep and AcMep were selected for further characterization to investigate additional features of keratinases in the M36 protease family.

Optimization of the reaction conditions for new family M36 keratinases

Preliminary experiments showed that at pH values below 6 or temperatures above 55 °C, the keratinases had very low activity (data not shown). Thus, using a pH range of 8–10 and a temperature range 25–45 °C, a 3-level full factorial design was set up to identify the optimal pH-temperature reaction conditions of FoMep (benchmark), PpMep, and

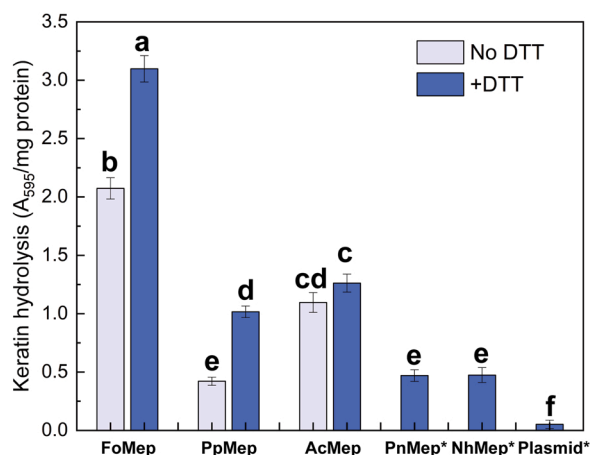


Fig. 3. The keratin hydrolysis ability of the M36 protease FoMep (from *F. oxysporum*), PpMep (from *P. pannorum*), AcMep (from *A. clavatus*), PnMep (from *P. nodorum*), NhMep (from *N. haematococca*) and empty plasmid (pPICZ α -A). The extent of keratin hydrolysis was tested in non-DTT pretreated keratin azure (light blue column) and in the presence of DTT (dark blue column; the keratin azure was pretreated by 2 mM DTT for 2 h). The reaction conditions were pH 7.5 (Tris-HCl buffer) at 30 °C for 1 h, *the product was tested after 24-h reaction with addition of DTT. ANOVA analysis at the 0.05 level. The absorbance of the total released azure was measured at 595 nm (A₅₉₅), and the data was reported as total A₅₉₅/mg protein. Data are given as average values of triplicate reactions and shown \pm the standard deviation; different letters a-f above columns indicate statistically significant differences in the keratinolytic hydrolysis ability of the enzymes in the keratinolytic test reactions ($p \leq 0.05$).

AcMep. Based on the experimental data, the keratinase activity response to reaction conditions could be established by response surface modeling. The predicted data model showed that for FoMep the reaction conditions for achieving maximum activity were pH 8.6 and 38 °C (Fig. 4a), for PpMep, pH 8.4 and 40 °C (Fig. 4b) and for AcMep pH 8.3 and 45 °C (Fig. 4c). Further, within the limits of pH 8.0–10.0, pH had a significant, negative effect on the activity of FoMep (Fig. 4a, Suppl. Table S2) and PpMep (Fig. 4b, Suppl. Table S2), but there was no significant effect of temperature. In contrast, for AcMep, activity was significantly affected, $p \leq 0.05$, by temperature and pH, i.e. activity increased at higher temperature and lower pH (Suppl. Table S2) (this is also evident from Fig. 4c). There was furthermore a strong interaction between these two factors implying that the activity was more affected by the temperature at low pH (Suppl. Table S2). The strength and reliability of the models was verified by the R^2 of FoMep, PpMep and AcMep of 0.977, 0.968 and 0.980, respectively, and the high Q^2 values (Suppl.

Table S2).

The predicted optimal conditions for FoMep, PpMep and AcMep activity agreed with the experimentally obtained activity values; the experimental values were 18.9, 7.6 and 9.1 U/mM, respectively (the predicted values were 17.0, 7.1 and 9.7 U/mM, respectively at the optimal reaction conditions). The temperature optimum of 38–45 °C for the three M36 keratinases was lower than that commonly reported for endo-acting keratinases in S1, S8 and M4 families (50–80 °C) [11]. However, all three enzymes had optimal pH values around 8.5, which is in accord with data for other reported keratinases, e.g. from S8 and M4 families [11,47].

Kinetics, thermal stability and specificity of the M36 keratinases

The kinetic parameters of FoMep, PpMep and AcMep were estimated on keratin azure. The presented kinetic curves, rate data versus varying substrate concentration were hyperbolic supporting that the enzyme reactions followed Michaelis-Menten kinetics (Suppl. Fig. S6). FoMep exhibited the highest catalytic rate with a k_{cat} of $\sim 0.8 \text{ min}^{-1}$ (calculated from a V_{max} of 5.4 $\mu\text{M}/\text{min}$) and lowest K_M of 143 mg/mL (Table 3); consequently FoMep also had the highest k_{cat}/K_M of $5.2 \times 10^{-3} \text{ min}^{-1} \text{ mg}^{-1} \text{ mL}$. The k_{cat} of PpMep and AcMep were 3–5 times lower, and the k_{cat}/K_M values for PpMep and AcMep were lower than that of FoMep.

The thermal stability and thus the half-life of PpMep were higher than those of the other enzymes at all temperatures tested (Table 3,

Table 3

Comparison of kinetic parameters, thermal stability and substrate specificity of the keratinases FoMep, PpMep and AcMep. Different superscript letters a, b, indicate significantly different parameter values among the enzymes ($p < 0.05$).

Item	FoMep	PpMep	AcMep
Kinetic parameters			
V_{max} ($\mu\text{M}/\text{min}$)	5.4 ^a \pm 0.7	2.4 ^b \pm 0.2	3.5 ^a \pm 0.5
K_M (mg keratin azure/mL)	143 ^b \pm 32.2	177 ^b \pm 21.8	249 ^a \pm 52.7
k_{cat} (min^{-1})	0.75	0.14	0.23
k_{cat}/K_M ($\times 10^{-3} \text{ min}^{-1} \text{ mg}^{-1} \text{ mL}$)	5.2	0.8	0.9
Thermal stability			
Thermal inactivation rate constant, k_D (min^{-1})			
30 °C	0.0083	0.0063	0.0068
40 °C	0.0179	0.0091	0.0175
50 °C	0.0827	0.0676	0.1588
60 °C	0.4936	0.2728	0.5112
Half-life (min)			
30 °C	83.5	110.0	101.9
40 °C	38.7	76.2	39.6
50 °C	8.4	10.3	4.4
60 °C	1.4	2.5	1.4
Substrate specificity			
Azocasein (U/mM)	3442	1040	802
Keratin azure (U/mM)	18.9	7.6	9.1
Azokeratin (U/mM)	1162	421	379

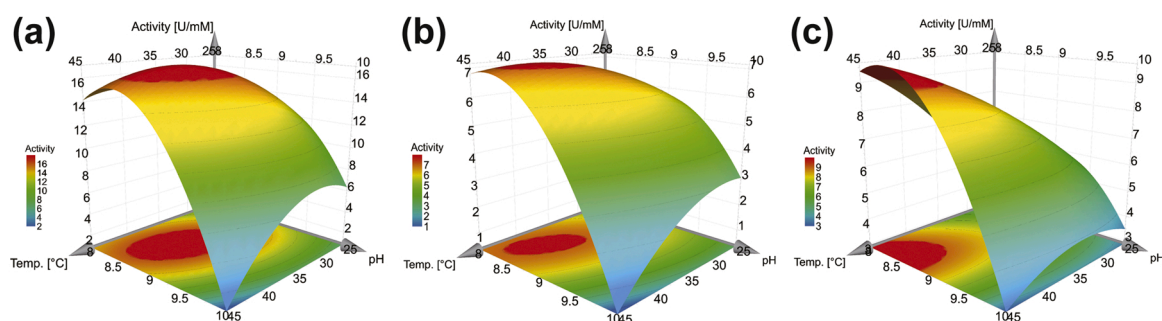


Fig. 4. Response surface plots of activity on keratin azure catalyzed by three keratinases as a function of pH and temperature (°C). (a) FoMep; (b) PpMep; (c) AcMep.

Suppl. Fig. S7). Although a half-life of 4–10 min at e.g. 50 °C is indicative of an enzyme which is useable at elevated temperatures, all known M36 keratinases, including now FoMep, PpMep, and AcMep, have lower thermal stability than other serine keratinases [7,48]. In view of their industrial production and applications, the engineering of M36 keratinases for enhanced thermal stability is worth investigating.

When examining substrate specificity, FoMep, PpMep and AcMep exhibited greatest activity on azocasein, with activities of 3442, 1040 and 802 U/mM, respectively, and as expected this “general proteolytic activity” was higher than the activity of the enzymes on keratin (Table 3). FoMep was the best enzyme among the three, but AcMep was more active than PpMep on keratin azure (sheep wool), whereas PpMep was more active than AcMep on azokeratin (pig bristle and hooves) (Table 3).

Structural modeling and catalysis theory analysis

Structural models of the catalytic domains of PpMep and AcMep were generated based on AfuMep [27]. PpMep and AfuMep had 67 % sequence identity, whilst AcMep and AfuMep had 78 % sequence identity (as examined using HHpred) (Fig. 5). The catalytic domains of PpMep, AcMep and AfuMep thus appeared to be similar, all of them

including 8 main α -helices and 6 β -strands (Fig. 5). By ProQ, the PpMep and AcMep models were judged as “extremely good” (LGscore 6.6, 4.7, respectively) [28]; there was only one outlier (T33) for PpMep and AcMep structures in the Ramachandran plot and 98.5 % (PpMep) and 99.4 % (AcMep) of the residues were in the favored region shown in the prediction server Zlab [29], indicating that the PpMep and AcMep homology models were reliable (data not shown). The pro-domain of AfuMep consists a tandem repeat of cystatin-like folds whose C-terminal end is buried in the active-site cleft of the catalytic domain (Fig. 5a, purple) [27]. The auto-proteolytic activation of the enzyme occurs early during its expression in the host, therefore the mature M36 AfuMep could have proteolytic activity [27]. In all three enzymes, it appears that the catalytic zinc ion (Zn^{2+}) is situated at the bottom of a long cleft in the middle of the catalytic domain (Fig. 5b). The structure also clearly shows how the Zn^{2+} is tetrahedrally bound by two histidines (H184, H188 both within the conserved motif, Fig. 2) and the glutamate moiety in the motif (E185) and additionally coordinated by an additional glutamate (E214) (Fig. 5c,d). The zinc ligands (H, H, and E) may thus stabilize the Zn^{2+} via hydrogen bonds, as seen in other proteases [49], and this tetrahedron moiety thus provides a key part of the mechanism by which PpMep and AcMep may act on keratin (Fig. 5d): water plays the role of nucleophile and the nucleophilicity of the water molecule could be

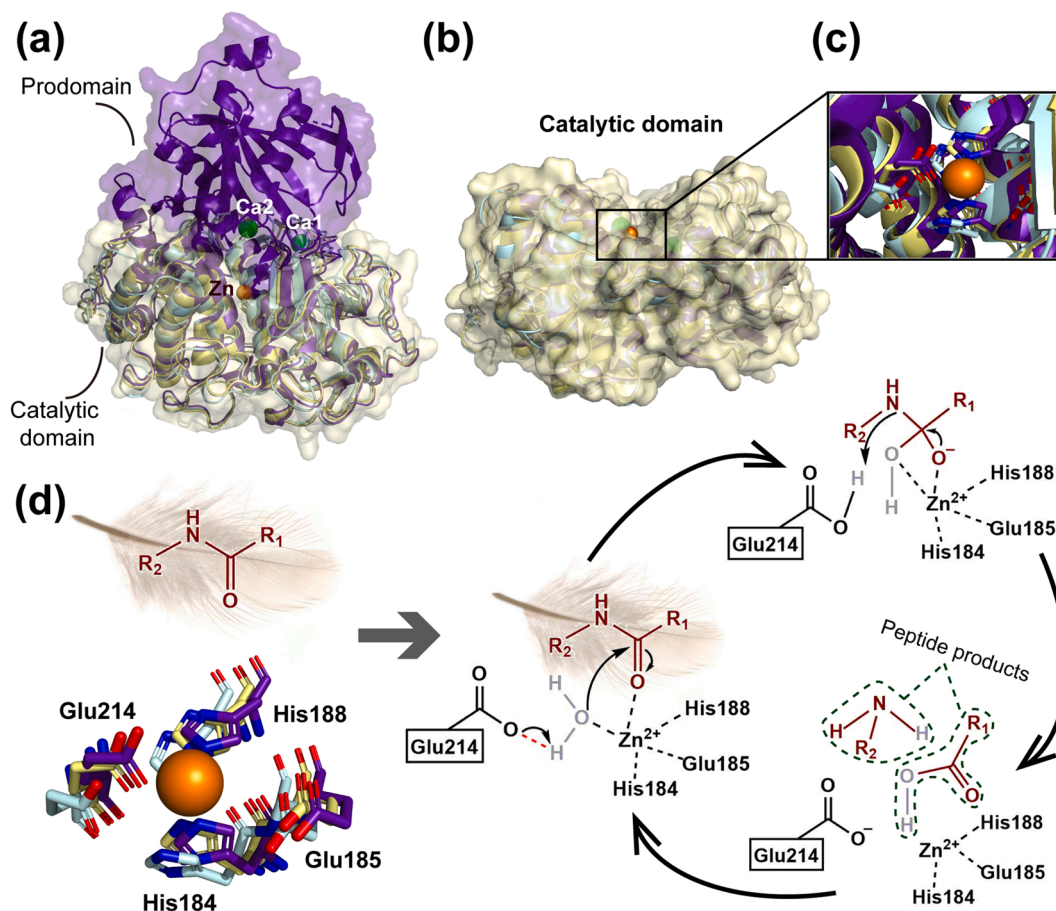


Fig. 5. Homology modeled structures of PpMep (keratinase from *P. pannorum*), AcMep (keratinase from *A. clavatus*) and a display of the presumed catalytic mechanism of PpMep and AcMep (model structure is based on the crystal structure of M36 protease AfuMep derived from *A. fumigatus* (PDB ID 4K90; MER0001400). (a) Purple: Crystal structure of M36 protease AfuMep; Blue: Homology model of PpMep, Yellow: Homology model of AcMep; The Ca1 and Ca2 (spheres in green), zinc ions (spheres in orange) of AfuMep are also shown. (b) The surface of the AfuMep catalytic domain. The active site is at the bottom of the cleft near the zinc ion (orange sphere). (c) The active sites of PpMep, AcMep and AfuMep proteases (overlay) are displayed as sticks (H184, E185, H188, E214); numbers correspond to the alignment snip in Fig. 2. (d) The hydrolytic mechanism of PpMep and AcMep: deprotonation of the nucleophilic water molecule; nucleophilic attack of the resulting hydroxide on the scissile carbonyl carbon; protonation of the scissile amide group, followed by cleavage of the peptide bond.

enhanced by having both the hydrogen protons bound to E214 while at the same time having the oxygen ligand bound to the Zn^{2+} (Fig. 5d). These tripartite interactions would leave the remaining lone pair directed toward the carbonyl carbon of the substrate and aligned for nucleophilic attack (Fig. 5d). Furthermore, proton transfer from E214 to the nitrogen atom facilitates breakdown of the tetrahedral intermediate and subsequent product formation (Fig. 5d) [50].

Conclusions

The members of MEROPS peptidase family M36 were classified into 11 groups based on conserved unique peptide patterns using the bioinformatics tool CUPP. Combined analysis of CUPP grouping and phylogenetic clustering of these M36 keratinases and peptidases revealed that members in CUPP group 1 have a high possibility of being keratinolytic proteases: group 1 contained all the confirmed M36 keratinases and an additional 71 peptidases of unknown keratinolytic function. Group 1 members are derived from diverse fungi in four different taxonomical classes, namely *Eurotiomycetes*, *Sordariomycetes*, *Dothideomycetes* and *Leotiomycetes*. One keratinase (FoMep) and four peptidases of unknown function (PpMep, AcMep, PnMep and NhMep derived from four organisms in different taxonomical classes) from group 1 were selected for heterologous expression in *P. pastoris* followed by keratinolytic activity testing. The results showed that all the selected M36 proteases had keratinolytic ability on keratin azure (particularly when DTT was added prior to reaction). FoMep, AcMep and PpMep were able to catalyze the hydrolysis of keratin, even without DTT pretreatment. The results confirm the relevance of CUPP grouping to function of proteases. To learn more about M36 keratinases, FoMep, AcMep and PpMep were further characterized. The optimal conditions for catalytic activity of FoMep, PpMep and AcMep were pH 8.6 at 38 °C, 40 °C and pH 8.4, and pH 8.3 at 45 °C, respectively. The kinetics analysis revealed that FoMep performed best, with a k_{cat} of 0.8 min^{-1} , and a k_{cat}/K_M of $5.24 \times 10^{-3} \text{ min}^{-1} \text{ mg}^{-1} \text{ mL}$. As regards thermal stability, PpMep had a longer half-life over the range 30–60 °C than FoMep and AcMep, although all three had half-lives of several minutes at 50 °C and even at 60 °C, none of the three were highly thermostable. PpMep displayed higher specificities than AcMep on pig bristles and hooves azokeratin, whereas AcMep performed better than PpMep when catalyzing the degradation of sheep wool, i.e. keratin azure. In summary, the current study presents a new application of CUPP for classifying proteases and the usefulness of CUPP grouping as a tool for discovering new keratinases. The approach identified 4 new fungal keratinases, of which notably AcMep and PpMep were expressed successfully in *P. pastoris* and were verified to have keratinolytic activity. This successful application of the CUPP concept provides a reference for discovering new functional proteases via CUPP bioinformatics.

Author contributions

All authors designed the study. JWQ conducted all experiments and analyzed the results. All the authors especially KB and JWQ carried out the CUPP analysis. CW designed the protease expression related combined plasmid and helped with the experiment. KB, CW, and ASM supervised the work. JWQ and ASM wrote the manuscript. All authors reviewed and approved the manuscript.

Declaration of Competing Interest

The authors declare that they have no conflicts of interest.

Acknowledgement

This work was supported by a China Scholarship Council (CSC) grant #201806910049 (to JWQ) and the Technical University of Denmark.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.nbt.2022.01.004>.

References

- [1] Daroit DJ, Brandelli A. A current assessment on the production of bacterial keratinases. *Crit Rev Biotechnol* 2014;34:372–84. <https://doi.org/10.3109/07388551.2013.794768>.
- [2] Sharma R, Devi S. Versatility and commercial status of microbial keratinases: a review. *Rev Environ Sci Biotechnol* 2018;17:19–45. <https://doi.org/10.1007/s11157-017-9454-x>.
- [3] Srivastava B, Khatri M, Singh G, Arya SK. Microbial keratinases: an overview of biochemical characterization and its eco-friendly approach for industrial applications. *J Clean Prod* 2020;252:119847. <https://doi.org/10.1016/j.jclepro.2019.119847>.
- [4] Jin H-S, Song K, Baek J-H, Lee J-E, Kim DJ, Nam G-W, et al. Identification of matrix metalloproteinase-1-suppressive peptides in feather keratin hydrolysate. *J Agric Food Chem* 2018;66:12719–29. <https://doi.org/10.1021/acs.jafc.8b05213>.
- [5] Li Z, Reimer C, Picard M, Mohanty AK, Misra M. Characterization of chicken feather biocarbon for use in sustainable biocomposites. *Front Mater* 2020;7:1–12. <https://doi.org/10.3389/fmats.2020.00003>.
- [6] Sharma S, Gupta A. Sustainable management of keratin waste biomass: applications and future perspectives. *Braz Arch Biol Technol* 2016;59:1–14. <https://doi.org/10.1590/1678-4324-2016150684>.
- [7] Benkiar A, Nadia ZJ, Badis A, Rebzani F, Soraya BT, Reikik H, et al. Biochemical and molecular characterization of a thermo- and detergent-stable alkaline serine keratinolytic protease from *Bacillus circulans* strain DZ100 for detergent formulations and feather-biodegradation process. *Int Biodeterior Biodegrad* 2013; 83:129–38. <https://doi.org/10.1016/j.ibiod.2013.05.014>.
- [8] Valencia R, González V, Undabarrena A, Zamora-Leiva L, Ugalde JA, Cámara B. An integrative bioinformatic analysis for keratinase detection in marine-derived *Streptomyces*. *Mar Drugs* 2021;19:286. <https://doi.org/10.3390/md19060286>.
- [9] Liang S, Deng J-J, Zhang M-S, Luo Z-Y, Lu W-J, Lu D-L, et al. Promotion of feather waste recycling by enhancing reducing power and keratinase activity of *Streptomyces* sp. SCUT-3. *Green Chem* 2021;23:5166–78. <https://doi.org/10.1039/d1gc00677k>.
- [10] Huang Y, Busk PK, Herbst FA, Lange L. Genome and secretome analyses provide insights into keratin decomposition by novel proteases from the non-pathogenic fungus *Oryzopsis korvina*. *Appl Microbiol Biotechnol* 2015;99:9635–49. <https://doi.org/10.1007/s00253-015-6805-9>.
- [11] Qiu J, Wilkens C, Barrett K, Meyer AS. Microbial enzymes catalyzing keratin degradation: classification, structure, function. *Biotechnol Adv* 2020;44:107607. <https://doi.org/10.1016/j.biotechadv.2020.107607>.
- [12] Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res* 2018;46: D624–32. <https://doi.org/10.1093/nar/gkx1134>.
- [13] Rawlings ND, Barrett AJ. Introduction: metallopeptidases and their clans. In: Barrett AJ, Rawlings ND, Fred Woessner J, editors. *Handbook of proteolytic enzymes*. Academic Press; 2004. p. 231–67. <https://doi.org/10.1016/B978-0-12-079611-3.50075-6>.
- [14] Rawlings ND, Barrett AJ. Introduction: serine peptidases and their clans. In: Rawlings ND, Salvesen G, editors. *Handbook of proteolytic enzymes*. Academic Press; 2013. p. 2491–523. <https://doi.org/10.1016/B978-0-12-382219-2.00559-7>.
- [15] Kang E, Jin HS, La JW, Sung JY, Park SY, Kim WC, et al. Identification of keratinases from *Fervidobacterium islandicum* AW-1 using dynamic gene expression profiling. *Microb Biotechnol* 2020;13:442–57. <https://doi.org/10.1111/1751-7915.13493>.
- [16] Chaya E, Suzuki T, Karita S, Hanya A, Yoshino-Yasuda S, Kitamoto N. Sequence analysis and heterologous expression of the wool cuticle-degrading enzyme encoding genes in *Fusarium oxysporum* 26-1. *J Biosci Bioeng* 2014;117:711–4. <https://doi.org/10.1016/j.jbiosc.2013.11.012>.
- [17] Brouta F, Descamps F, Monod M, Vermout S, Losson B, Mignon B. Secreted metalloprotease gene family of *Microsporium canis*. *Infect Immun* 2002;70:5676–83. <https://doi.org/10.1128/iai.70.10.5676-5683.2002>.
- [18] Banerjee A, Sahoo DK, Thatoi H, Pati BR, Mondal KC, Sen A, et al. Structural characterization and active site prediction of bacterial keratinase through molecular docking. *J Bioinform* 2014;1:67–82.
- [19] Wu WL, Chen MY, Tu IF, Lin YC, Eswarkumar N, Chen MY, et al. The discovery of novel heat-stable keratinases from *Meiothermus taiwanensis* WR-220 and other extremophiles. *Sci Rep* 2017;7:1–12. <https://doi.org/10.1038/s41598-017-04723-4>.
- [20] Wang L-S, Leebens-Mack J, Wall PK, Beckmann K, DePamphilis CW, Warnow T. The impact of multiple protein sequence alignment on phylogenetic estimation. *IEEE/ACM Trans Comput Biol Bioinforma* 2009;8:1108–19. <https://doi.org/10.1109/tcbb.2009.68>.
- [21] Barrett K, Hunt CJ, Lange L, Meyer AS. Conserved unique peptide patterns (CUPP) online platform: peptide-based functional annotation of carbohydrate active enzymes. *Nucleic Acids Res* 2020;48:W110–5. <https://doi.org/10.1093/nar/gkaa375>.

- [22] Barrett K, Lange L. Peptide-based functional annotation of carbohydrate-active enzymes by conserved unique peptide patterns (CUPP). *Biotechnol Biofuels* 2019; 12:102. <https://doi.org/10.1186/s13068-019-1436-5>.
- [23] Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 2012;28:3150–2. <https://doi.org/10.1093/bioinformatics/bts565>.
- [24] Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE)*; 2010. p. 1–8.
- [25] Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 2019;47:W256–9. <https://doi.org/10.1093/nar/gkz239>.
- [26] Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 2005;33:244–8. <https://doi.org/10.1093/nar/gki408>.
- [27] Fernández D, Russi S, Vendrell J, Monod M, Pallarès I. A functional and structural study of the major metalloprotease secreted by the pathogenic fungus *Aspergillus fumigatus*. *Acta Crystallogr Sect D Biol Crystallogr* 2013;69:1946–57. <https://doi.org/10.1107/S0907444913017642>.
- [28] Wallner B, Elofsson A. Can correct protein models be identified? *Protein Sci* 2003; 12:1073–86. <https://doi.org/10.1110/ps.0236803>.
- [29] Anderson RJ, Weng Z, Campbell RK, Jiang X. Main-chain conformational tendencies of amino acids. *Proteins Struct Funct Bioinforma* 2005;60:679–89. <https://doi.org/10.1002/prot.20530>.
- [30] Armenteros JJA, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 2019;37:420–3. <https://doi.org/10.1038/s41587-019-0036-z>.
- [31] Silva IR, Larsen DM, Meyer AS, Mikkelsen JD. Identification, expression, and characterization of a novel bacterial RGI Lyase enzyme for the production of bio-functional fibers. *Enzyme Microb Technol* 2011;49:160–6. <https://doi.org/10.1016/j.enzmictec.2011.04.015>.
- [32] Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, et al. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* 1999;112:531–52. <https://doi.org/10.1385/1-59259-584-7:531>.
- [33] Yang X, Naughton SX, Han Z, He M, Zheng YG, Terry Jr AV, et al. Mass spectrometric quantitation of tubulin acetylation from pepsin-digested rat brain tissue using a novel stable-isotope standard and capture by anti-peptide antibody (SISCAPA) method. *Anal Chem* 2018;90:2155–63. <https://doi.org/10.1021/acs.analchem.7b04484>.
- [34] Zipfel E, Grezes JR, Seiffert W, Zimmermann HW. Romanowsky dyes and romanowsky-Giemsa effect. 1. Azure B, purity and content of dye samples, association (author's transl). *Histochemistry* 1981;72:279–90. <https://doi.org/10.1007/BF00517141>.
- [35] Gonzalo M, Espersen R, Al-Soud WA, Cristiano Falco F, Hägglund P, Sørensen SJ, et al. Azo dyeing of α -keratin material improves microbial keratinase screening and standardization. *Microb Biotechnol* 2020;13:984–96. <https://doi.org/10.1111/1751-7915.13541>.
- [36] Huang Y, Łężyk M, Herbst FA, Busk PK, Lange L. Novel keratinolytic enzymes, discovered from a talented and efficient bacterial keratin degrader. *Sci Rep* 2020; 10:1–11. <https://doi.org/10.1038/s41598-020-66792-2>.
- [37] Braissant O, Keiser J, Meister I, Bachmann A, Wirz D, Göpfert B, et al. Isothermal microcalorimetry accurately detects bacteria, tumorous microtissues, and parasitic worms in a label-free well-plate assay. *Biotechnol J* 2015;10:460–8. <https://doi.org/10.1002/biot.201400494>.
- [38] Abe M, Ohsato M, Kawamura T, Ogino K. Interaction between anionic surfactants and oil dyes in the aqueous solutions. V. The effects of the oxyethylene chain length and/or the alkyl chain length in surfactant molecule on the tautomerism of an azo dye. *J Colloid Interface Sci* 1985;104:228–33. [https://doi.org/10.1016/0021-9797\(85\)90027-X](https://doi.org/10.1016/0021-9797(85)90027-X).
- [39] Fedorova ND, Khaldi N, Joardar VS, Maiti R, Amedeo P, Anderson MJ, et al. Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. *PLoS Genet* 2008;4:e1000046. <https://doi.org/10.1371/journal.pgen.1000046>.
- [40] Lopes FC, Silva LADE, Tichota DM, Daroit DJ, Velho RV, Pereira JQ, et al. Production of proteolytic enzymes by a keratin-degrading *Aspergillus niger*. *Enzyme Res* 2011;2011:1–9. <https://doi.org/10.1017/S0079497X00005089>.
- [41] Noronha EF, De Lima BD, De Sá CM, Felix CR. Heterologous production of *Aspergillus fumigatus* keratinase in *Pichia pastoris*. *World J Microbiol Biotechnol* 2002;18:563–8. <https://doi.org/10.1023/A:1016341702908>.
- [42] Hane JK, Lowe RGT, Solomon PS, Tan K-C, Schoch CL, Spatafora JW, et al. Dothideomycete–plant interactions illuminated by genome sequencing and EST analysis of the wheat pathogen *Stagonospora nodorum*. *Plant Cell* 2007;19:3347–68. <https://doi.org/10.1105/tpc.107.052829>.
- [43] Sarker S, Das S, Lavers JL, Hutton I, Helbig K, Imbery J, et al. Genomic characterization of two novel pathogenic avipoxviruses isolated from pacific shearwaters (*Ardenna* spp.). *BMC Genomics* 2017;18:1–26. <https://doi.org/10.1186/s12864-017-3680-z>.
- [44] Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, et al. The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. *PLoS Genet* 2009;5:e1000618. <https://doi.org/10.1371/journal.pgen.1000618>.
- [45] Ke Y, Yuan XM, Li JS, Zhou W, Huang XH, Wang T. High-level expression, purification, and enzymatic characterization of a recombinant *Aspergillus sojae* alkaline protease in *Pichia pastoris*. *Protein Expr Purif* 2018;148:24–9. <https://doi.org/10.1016/j.pep.2018.03.009>.
- [46] Espersen R, Falco FC, Hägglund P, Gernaey KV, Lantz AE, Svensson B. Two novel S1 peptidases from *Amycolatopsis keratinophila* subsp. *keratinophila* D2^T degrading keratinous slaughterhouse by-products. *Appl Microbiol Biotechnol* 2020;104: 2513–22. <https://doi.org/10.1007/s00253-020-10380-x>.
- [47] Nnolim NE, Udenigwe CC, Okoh AI, Nwodo UU. Microbial keratinase: next generation green catalyst and prospective applications. *Front Microbiol* 2020;11: 580164. <https://doi.org/10.3389/fmicb.2020.580164>.
- [48] Bouacem K, Bouanane-Darenfed A, Zaraï Jaouadi N, Joseph M, Hacene H, Ollivier B, et al. Novel serine keratinase from *Caldicoprobacter algeriensis* exhibiting outstanding hide dehairing abilities. *Int J Biol Macromol* 2016;86:321–8. <https://doi.org/10.1016/j.ijbiomac.2016.01.074>.
- [49] Fukasawa KM, Hata T, Ono Y, Hirose J. Metal preferences of zinc-binding motif on metalloproteases. *J Amino Acids* 2011;2011:574816. <https://doi.org/10.4061/2011/574816>.
- [50] Hernick M, Fierke C. Mechanisms of metal-dependent hydrolases in metabolism. *Compr Nat Prod II Chem Biol* 2010;8:547–81. <https://doi.org/10.1016/b978-008045382-8.00178-7>.