Molecular Cloning, Expression and Characterisation of a bacterial myrosinase from *Citrobacter* Wye1

Fatma Cebeci [a,1,*], Melinda J. Mayer [b], John T. Rossiter [c], Richard Mithen [a,2] and Arjan Narbad [a,b]

[a] Food Innovation and Health Institute Strategic Programme, Quadram Institute Bioscience, Norwich, UK
[b]Gut Microbes and Health Institute Strategic Programme, Quadram Institute Bioscience, Norwich, UK
[c]Department of Life Sciences, Imperial College London, London, UK
[1]Department of Nutrition and Dietetics, Bayburt University, Bayburt, Turkey
[2]Liggins Institute, The University of Auckland, Auckland, New Zealand
*Correspondence: fatmacebeci@bayburt.edu.tr; ORCID: 0000-0003-4715-6689

Abstract: Glucosinolates are plant natural products which on degradation by myrosinases give rise to the beneficial bioactive isothiocyanates. Recently, a myrosinase activity was detected in a *Citrobacter* strain isolated from soil. This enzyme was purified enabling its amino acid sequence and gene sequence (*cmyr*) to be determined. In order to study this myrosinase it was necessary to establish an expression system that would enable future work such as a structural determination of the protein to be carried out.

The myrosinase gene was amplified, cloned and expressed in *Escherichia coli* with a 6XHis-tag. The heterologous expression of *cmyr* enabled relatively large amounts of myrosinase to be produced (3.4 mg cmyr/ 100 ml culture). Myrosinase activity was determined by mixing substrate and enzyme and determining glucose release. Optimum pH and temperature were determined to be pH 6.0 and 25°C for the Ni-NTA purified protein. The kinetic parameters of the purified myrosinase were determined using sinigrin as a substrate. K_m and V_{max} were estimated as 0.18 mM and 0.033 mmol/min/mg respectively for sinigrin under optimum conditions and compared to other kinetic data for myrosinases. The substrate specificity of myrosinase was determined having the highest affinity for sinigrin followed by glucoiberin, progoitrin, glucoerucin, glucoraphanin and glucotropaeolin.

Keywords: *Citrobacter*; myrosinase; glucosinolate; biotransformation

1. Introduction

Glucosinolates are secondary metabolites found in the Brassicaceae such as broccoli (*Brassica oleracea* var. *italic*), cabbage (*Brassica oleracea* var. *capitata* f. *alba*) and cauliflower (*Brassica oleracea* var. *botrytis*). The importance of glucosinolates lies in their enzymatic degradation products [1], [2] notably the isothiocyanates which have health promoting effects [3][4]. They also have a key role in plant insect interactions as well as in allopathy [5]–[10]. Glucosinolates are hydrolysed by myrosinases (EC 3.2.1.147, thioglucosidase glucohydrolase) on tissue damage to the plant to form isothiocyanates (ITCs) (Figure 1) unless specifier proteins are present in which case hydrolysis can lead to the formation of nitriles, thiocyanates and epithionitriles [11]. ITCs are known to have antioxidant, immunostimulatory, anti-inflammatory, antiviral and antibacterial properties as well as chemopreventive and therapeutic agents [4], [12], [13].

The structure of plant myrosinase was revealed some years ago when the crystal structure of myrosinase from *Sinapis alba* seeds was elucidated. The *S. alba* myrosinase utilizes a nucleophile (glutamate) and ascorbate as a catalytic base. It is a dimer stabilized by a zinc atom and has a (β/α)8-barrel structure similar to the GH1 (glucosidase family 1) enzymes [14], [15].

The presence of myrosinases extends beyond plants and are found in aphids [16]–[20], beetles [21], bacteria [22]–[25] and fungi [26]–[28] and are likely to occur in many diverse organisms. In comparison to plants less is known about the bacterial myrosinases [29]. The first bacterial myrosinase was purified in 1974 from *Enterobacter cloacae* but work was not extended to investigate the amino acid sequence of this protein [22]. Later studies examined a range of bacteria for their role in glucosinolate metabolism as well as the identification and characterization of putative myrosinases [25], [30]. Several studies have described glucosinolate metabolism by human gut bacteria in rat models [31], [32] or using *in vitro* fermentation models [25], [30].

More recently a myrosinase from a *Citrobacter* strain isolated from soil was identified [33] while a recent study reported the presence of an operon responsible for glucosinolate metabolism in a gut bacterium *Bacteriodes thetaiotaomicron* [34].

Citrobacter Wye1 was isolated from soil using an enrichment method with sinigrin as the sole carbon source. Crude cell free extracts showed an active myrosinase which produced isothiocyanates. Using a combination of ion exchange and gel filtration chromatography, a myrosinase was purified to homogeneity and the N terminal amino acid sequence determined. Additional sequence determination enabled the gene to be identified from the sequenced genome. An enzyme kinetic study was also carried out on the crude protein extracts of *Citrobacter* Wye1 giving apparent K_m and V_{max} values. It was revealed that this myrosinase is a periplasmic β-glucosidase from the GH3 family with a signal peptide [33]. So far, identified myrosinases in plants, aphids or beetles fall into the GH1 family while *Citrobacter* Wye1 myrosinase (cmyr) belongs to the GH3 family. Plant myrosinases are activated by ascorbate to varying degrees [35] acting as a catalytic base [15] within the active site of the enzyme [36]. Although the myrosinase of *Citrobacter* strain from soil was found to be slightly activated by ascorbic acid [33], non-plant myrosinases can be neutral to the presence of ascorbic acid or inhibited by it [36].

The cmyr is not an abundant protein and requires relatively large scale fermentation and large amounts of the glucosinolate sinigrin making characterization studies prohibitively expensive. Thus, to enable characterization of this enzyme we have set out to overexpress *cmyr* in *Escherichia coli* and carry out a kinetic study with a variety of glucosinolate substrates.

2. Materials and Methods

2.1. Bacterial Strains

E. coli DH5α and *E. coli* (BL21DE3) were purchased from Thermo Fisher Scientific.

2.2.Multiple Alignment of Citrobacter Wye1 Myrosinase

Amino acid sequences of different myrosinases from different origins were selected using the Uniprot database. These myrosinases were aligned with cmyr using Clustal Omega [37].

2.3. Cloning and Expression of Myrosinase Gene

A cloning vector which has 6XHis-tag site in both N-terminal and C-terminal site (pET28b) was used as cmyr was reported to have a signal peptide. Forward and reverse primers were designed to introduce new restriction sites by changing nucleotides and to enable in-frame translation of C-terminal 6XHis-tag by insertion of a new nucleotide to the reverse primer. The primers used to amplify the *cmyr* and PCR conditions are given in Table 1 with the restriction sites underlined. The PCR insert was amplified using Phusion polymerase (Finnzymes) and double restricted by NdeI/XhoI and ligated to NdeI/XhoI double restricted pET28b vector (Novagen) then *E. coli* DH5α cells were transformed with this gene construct and the cells were plated on L agar with kanamycin (30 μg/ml final concentration).

Colonies were screened by colony PCR using primers T7P2 (5′-TGAGCGGATAACAATTCCC) and T7T (5′-GCTAGTTATTGCTCAGCGG) to select the positive transformants. The clones that gave a positive result for PCR were selected and inoculated in L Broth with kanamycin (30 μg/ml final concentration) and grown overnight at 37°C, with shaking at 250 rpm. The recombinant plasmids were extracted using EZNA Plasmid Mini Kit II (Omega Bio-Tek) and sequenced, then *E. coli* BL21 (DE3) was transformed with the recombinant plasmid for protein expression.

E. coli BL21 (DE3) expressing pET28b-*cmyr* was induced with 0.5 mM IPTG. Protein expression induction was carried out at 25°C for 24 h or at 37°C for 3 h. After protein induction, cell-free extracts (CFEs) were prepared by sonication (Soniprep 150, MSE) (sonicating for 15 s then on ice for 30 s, 7X). After centrifuged for 25 min, 13000 x g at 4°C, the supernatants were transferred into new clean tubes (soluble extracts), 500 μl resuspension buffer was added to the cell pellets (insoluble extracts). Soluble and insoluble fractions were run on SDS-PAGE (Invitrogen).

2.4. Purification of Myrosinase

For Western blot analysis, proteins were transferred onto a polyvinylidene difluoride membrane (Invitrogen) as manufacturer's instructions. His-tagged proteins were detected using an anti-His tag monoclonal antibody (Novagen) with alkaline phosphatase-linked anti-mouse immunoglobulin G (Sigma-Aldrich) as the secondary antibody and colorimetric detection with Sigma Fast BCIP (5-bromo-

4-chloro-3-indolylphosphate)-nitroblue tetrazolium as the substrate. Following confirmation by Western Blot, cmyr was purified by affinity purification using a Ni-NTA column. *E. coli* BL21 (DE3) cells with *cmyr* were grown in L broth with antibiotic (30 μg/ml final concentration) at 25°C for 4 h with 250 rpm shaking until an optical density (OD_{600}) of 0.6 was reached. The culture was induced with IPTG at a final concentration of 0.5 mM IPTG. The cells were harvested by centrifugation for 20 min, at 3200 xg 4°C. Ni-NTA agarose resin slurry (Qiagen) was used to perform purification of His-tagged proteins as recommended by the manufacturer, using an elution buffer of 10 mM Tris pH 8.0, 150 mM NaCl, 200 mM imidazole. Purified proteins were quantified using the Bradford assay (Bio-Rad) and visualized on 4-12% NuPage Bis-Tris gels in MES (2-(N-morpholino)ethanesulfonic acid) buffer stained with Simply Blue Safestain (Invitrogen).

Ni-NTA purified cmyr was dialysed to remove excessive amount of imidazole from Ni-NTA column purification. Spectra/Por porous membrane tubing with 500-1000Da cut off (Spectrum Labs) was used for dialysis. The protein was dialysed against 20 mM citrate phosphate buffer pH 6.0 for 18 h at 4°C. The protein concentration of the dialysed fractions was quantified by Bradford assay.

2.5 Measurement of Enzyme Activity

God-Perid assay was used to determine the enzyme activity of cmyr [38]. The absorbance readings at 420 nm were measured using 6715 UV/Vis Spectophotometer (JENWAY, UK). A calibration curve was prepared using glucose to quantify the glucose released due the enzyme activity. All activity assays were repeated 3 times using the same batch of enzyme (technical replicates).

2.6. Intact Mass Analysis

Dialysed protein sample was concentrated to 50 μM using Amicon Ultra 4-50 k filters (Millipore), this protein sample was analysed by LCMS to determine the intact mass of the protein. The analyses were carried out on a Synapt G2-Si mass spectrometer coupled to an Acquity UPLC system (Waters, UK). An aliquot of ~300 pmoles of protein was injected onto an Aeris Widepore 3.6u C4 Column 50 x 2.1 mm (Phenomenex). Elution was performed by a gradient of 5-90% acetonitrile in 5 min with a flow rate of 0.4 ml/min. Masslynx 4.1 software (Waters) was used to control the spectrometer and it was operated in positive MS-TOF and resolution mode with a capillary voltage of 2.5 kV and a cone voltage of 40 V. Leu-enkephalin peptide (1 ng/ml, Waters) was infused at 3 μl/min as a lock mass and measured every 20 s. The spectra were produced by combining a number of scans, and deconvoluted using the MaxEnt1 tool in Masslynx.

4.6. Characterisation of Myrosinase

The Ni-NTA purified, dialysed enzyme was used for further characterisation. Substrate specificity, the optimum pH and temperature conditions for myrosinase activity were investigated.Freshly prepared enzyme was used for enzyme kinetics experiments to determine K_m and V_{max} for sinigrin under optimum conditions.

The optimum pH was tested in 20 mM citrate phosphate buffer (pH range of 3.6-7.6). Dialysed cmyr was added to the 300 μl mixture of sinigrin (2 mM, final), protein (24 ng) in 20 mM citrate phosphate buffer at a pH range of 3.0 - 7.6. The assay mixture was incubated at 37°C for 1 h and reaction was stopped by boiling the tubes for 5 min to inactivate the myrosinase. God-Perid assay was used to determine the glucose released. The optimum temperature was tested over a range of temperatures (5-70°C). The reaction mixtures were set up in 300 μl of volume consisting sinigrin (2 mM), protein (24 ng) in citrate phosphate buffer pH 6.0, incubated for 30 min and glucose release was measured by God-Perid assay.

To assess the Michaelis-Menten constant (K_m) and maximum velocity (V_{max}) for sinigrin, 0.1-5 mM range of sinigrin was incubated with cmyr (24 ng) in 300 μl of citrate phosphate buffer pH 6.0 at 25°C for 10 min then activity was determined by God-Perid assay. The Michaelis Menten plot method was prepared to determine K_m and V_{max}. The myrosinase activity of cmyr was tested using 2 mM of different glucosinolates (sinigrin, glucoiberin, progoitrin, glucoerucin, gluroraphanin, glucotropaeolin, gluconasturtiin and glucobrassicin) by God-Perid assay. The assay was undertaken at optimum temperature and buffer conditions for myrosinase activity. In addition, the stability of myrosinase activity of cmyr was tested in 20 mM citrate phosphate buffer pH 6.0 over 5 weeks at 4°C.

3. Results

3.1. The Comparison of Citrobacter Wye1 Myrosinase to Known Myrosinases

The protein sequence of cmyr was aligned with other myrosinases identified from different sources and we confirmed that cmyr is quite distinct from other myrosinases. There is a high sequence similarity among plant myrosinases from thale cress (*Arabidopsis thaliana)*, wasabi (*Eutrema japonicum)*, white mustard (*S. alba*), rapeseed (*Brassica napus)* and Chineese kale (*B. oleracea* var. *alboglabra*). The nucleophile was reported to be Glu 409 residue in *S. alba* myrosinase and ascorbic acid substitutes the proton donor. On the other hand, Glu 374 and Glu 167 were reported to be the nucleophile and the proton donor respectively in cabbage aphid (*Brevicoryne brassicae*) myrosinase [18]. These two key amino acid residues were not conserved in cmyr and it was found to be quite distinct from aphid and plant myrosinases (See Supplementary, S1). The importance of the conserved motif 'SDW' (a feature of GH3 glucosidases) in cmyr possessing aspartic acid (D) as the nucleophile was suggested by Albaser et al. (2016) [33].

3.2. Cloning and Expression of Myrosinase Gene

The myrosinase gene was amplified by PCR and the gene sequence was submitted to Genbank (ALM58466.1). The gene was cloned into pET28b and *E. coli* DH5α cells were transformed with the recombinant plasmid. This recombinant plasmid was sequenced and confirmed to have no mutation in the gene insert. *E. coli* BL21(DE3) was transformed with the construct and the induction of protein expression was performed by addition of IPTG.

The soluble and insoluble cell free extracts (CFE) from protein expression trials were monitored by SDS-PAGE. Protein expression at 25°C for 24 h resulted in a good yield of expressed cmyr in the soluble fraction compared to protein expression at 37°C for 3 h, which resulted in in cmyr being predominantly in the insoluble fraction.

3.3. Western Blotting and Purification of Myrosinase

The CFEs were prepared and analysed for the presence of a 6X-His tag by Western blot analysis (Figure 2) using a His-tag antibody (Novagen). The incubation time for bacterial expression was reduced from 24 h to 4 h providing an increased protein yield.

The cmyr has two 6XHis-tags at N and C terminal site and the predicted size of the protein is ~75 kDa. However, on cleavage of the N-tagged then the predicted size is ~70 kDa with the one 6XHis-tag at C terminal site. The intact mass spectrum analysis showed that cmyr has a measured mass of 70,332 Da suggesting that the signal peptide was cleaved correctly (See Supplementary S2 for MS chromatogram).

Following expression, it was purified by Ni-NTA column (Figure 3) yielding approximately 3.4 mg cmyr/ 100 ml culture.

3.4. Characterisation of recombinant myrosinase

S. alba myrosinase (Sigma Aldrich) was used as a positive control and gave a positive result for myrosinase activity (0.33 ± 0.05 μmol glucose/min/mg enzyme).

The myrosinase activity of cmyr was tested in 20 mM citrate phosphate buffer within a pH range of 3.6-7.6, pH 6.0 was determined to be optimum for myrosinase activity (Figure 4). The enzyme activity was found to decrease beyond pH 6.0.

The optimum temperature was found to be 25°C under the conditions tested (Figure 5). The enzyme activity was reduced significantly beyond 30°C and further diminished over 50°C.

3.5. Enzyme Kinetics

The kinetic parameters of the Ni-NTA purified cmyr were determined using sinigrin as substrate. K_m and V_{max} were estimated as 0.18 mM and 0.033 mmol/min/mg for sinigrin at pH 6.0 and 25°C (Figure 6).

3.6. Substrate Specificity of cmyr

The cmyr showed activity towards all of the glucosinolates (Table 2) except gluconasturtiin and glucobrassicin. The specific activity of cmyr was highest for sinigrin, followed by glucoiberin > progoitrin > glucoerucin > glucoraphanin > glucotropaeolin. The cmyr had similar specific activity for glucoerucin, glucoiberin and progoitrin while that of glucoraphanin and glucotropaeolin was lower. Gluconasturtiin and glucobrassicin were not degraded by this enzyme under the conditions tested.

3.7. Activity of recombinant cmyr stored at 4°C

Dialysed cmyr was tested for activity in 20 mM citrate phosphate buffer pH 6.0 at 4°C. Results indicate that cmyr lost half of its original myrosinase activity within 6 days (Figure 7). At 5 weeks, 97% of the activity was lost.

4. Discussion

The main objective of this study was to develop an expression system for a bacterial myrosinase to enable sufficient amounts of protein for characterization studies. This was achieved using heterologous protein expression in *E. coli*. The recombinant protein was not stable and had a half-life of 6 days. This instability was also observed with the native protein [33] but not quantified.

There have been many K_m and V_{max} determinations on plant myrosinases [35], [36] while non plant K_m and V_{max} determinations have been limited to aphid myrosinases e.g. the cabbage aphid (*Brevicoryne brassicae*) myrosinase using sinigrin as a substrate and were found to be 0.41 mM (at pH 4.5, 30°C)[16] and 0.61 mM (at pH 5.5, 30°C)[17] for myrosinase purified by conventional chromatography and utilizing an expression system respectively.

In this study, we observed a lower K_m value (0.18 mM) for the cloned purified protein than that obtained by Albaser et al. (2016) (0.46 mM) where cell free protein extracts were used [33]. The K_m value obtained by Albaser et al. (2016) was an apparent K_m as a crude protein extract was used. Thus, other proteins in the extract may have influenced the outcome of the K_m determination. Also given the relatively short half-life of the myrosinase it is difficult to directly compare the K_m values for each of the protein preparations.

There have been several K_m determinations on plant myrosinase in the absence [39] and presence of ascorbate [40] where ascorbate acts as a catalytic base [15] and impacts on both K_m and V_{max} values. The necessity of activation of plant myrosinase by ascorbate most likely stems from its role in defense against pests and diseases where a rapid response is required. The myrosinase of the cabbage aphid is highly concentrated into microbodies within the thoracic muscles of the insect leading to rapid hydrolysis of the haemolymph glucosinolates and does not require a cofactor to augment activity [17], [41], [42]. The cmyr is an inducible enzyme and does not have a defense role but rather a nutritional one (glucose production from glucosinolate) and again there would be no real advantage in augmentation by a cofactor. Also ascorbate levels are unlikely to be of significant levels in the environment of this bacteria. It has been shown that a myrosinase from *Lepidium latifolium* L. is redox regulated [39] but that was not considered with respect to the cmyr.

The optimum temperature for myrosinase activity was found to be 25°C which is consistent with the native protein [33], but lower than some myrosinases [39], [40], [43]. The pH optimum for cmyr is 6 and is similar to that of some plant myrosinases such as from rapeseed (*B. napus*) [40] and tall whitetop (*L. latifolium* L.) [39]. Although the cmyr is unstable over time the determinations of the temperature and pH optima determinations would not be affected as the assay points are relative to one another.

In comparison to other myrosinase pH optima, cmyr was lower than a fungus myrosinase [26] and higher than cabbage aphid myrosinase [17]. The activity of cmyr was found to gradually decrease above pH 6.0, which was not the case with rapeseed and horseradish myrosinases [40], [43]. Li et al. (2005) reported that optimum pH was 5.7 for horseradish (*A. rusticana*) myrosinase and myrosinase activity was reported to retain 80% of the maximum activity between pH 5.0-8.0 values [43]. We observed a significant decrease in myrosinase activity above optimum pH especially above pH 7.0 for cmyr. This might be explained by the origin of the *Citrobacter* Wye1 which was isolated from soil with neutral pH (UK Soil Observatory, 2021) [44]. Myrosinases are also likely to be structurally diverse as cmyr belongs to the GH3 family of glucosidases and not the GH1 family, as is the case with most myrosinases characterised so far.

This study investigated the enzyme activity towards additional substrates compared to the previous study [33]. The cmyr showed differing specificity for each glucosinolate tested. The preference of aliphatic glucosinolates (glucoiberin, progoitrin, glucoerucin and glucoraphanin) over aromatic glucosinolates (glucotropaeolin) might indicate the importance of the side chain for myrosinase activity of cmyr. As cmyr showed a lower specificity for glucosinolates with longer side chains such as glucoraphanin compared to glucoiberin, the length of the side chains seems to be an important determinant for the specificity of cmyr. Comparing specific activity for glucoerucin and glucoraphanin

suggests that the oxidation of sulfur in methionine derived side chain is important as the polar nature of the sulfinyl group may result in unfavourable interactions within the active site of the enzyme. It was also reported that *L. agilis* R16 does not metabolise glucoraphanin to a significant degree which suggests that the myrosinases are hindered by the steric effects of the methylsulfinyl group [25].

The presence of indole or aromatic rings in the side chain affects specificity since cmyr shows very low or no activity towards glucotropaeolin, gluconasturtiin and glucobrassicin (Table 2). The specificity of myrosinases can vary quite markedly [16], [17], [39], [45], [46]. For example, as with cymr the cabbage aphid myrosinase was found to have higher specificity for sinigrin than glucotropaeolin in two separate studies [16], [17]. Even though previous work reported formation of ITC from gluconasturtiin [33], our study did not show any activity for this substrate. A potential reason for this might be the effect of 6X-His tag attached to the protein at C-terminal site. There are studies reporting the detrimental effects of a polyhistidine tag at C-terminal site on enzyme stability and activity [47], [48] and changes in substrate specificity [49].

The importance of a conserved amino acid sequence, SDW, was reported for the enzyme activity of cmyr and aspartic acid (D) was suggested as the catalytic nucleophile by Albaser et al. (2016) [33]. The mechanism of glucosinolate hydrolysis by cmyr is likely to follow acid/base catalysis together with a catalytic nucleophile. A putative mechanism has been suggested by Albaser et al. (2016) [33] and is shown in Figure 8. Crystallography studies, however, are required to identify the active site of the enzyme which would enable modification strategies to increase activity as well broadening substrate specificity. Myrosinase stabilization studies together with a view to obtaining a crystal structure will go some way to elucidating the mechanism of cmyr.

Whilst a great deal of sequence information is available for plant myrosinases relatively little sequence data is available for bacterial myrosinases [29]. Multiple sequence alignments were performed to compare cmyr with identified myrosinases against the UniProt database but there are no new myrosinase sequences having a significant match to cmyr to date. To enable a better understanding of the metabolism of glucosinolates in the human gut/soil it is important to identify bacteria that can metabolise glucosinolates as well as the putative myrosinases that carry out the hydrolysis. Previous work with *E. cloacae* showed the presence of a myrosinase but there is no sequence data for this activity. There is, however, 70% homology of a *E. cloacae* protein with the cmyr that shares the characteristic SDW signature of the GH3 β-*O*-glucosidases [33] which is tempting to speculate that this is the myrosinase of *E. cloacae*. Elucidation of myrosinase bacterial sequences will eventually allow more meaningful studies to be carried out with gut models. Although cmyr is of soil origin, it is known that the human gut can also harbor *Citrobacter* species. For example, one of the human gut isolates from our laboratory showed glucosinolate degrading ability (unpublished data) and 16S rDNA and genome sequencing revealed that it was a strain of *Citrobacter freundii* [50].

Mustard crops that have high concentrations of sinigrin are used as biofumigants in a process known as green manuring for pest and disease control. These products produce high levels of isothiocyanates but one potential issue here is that soil microbiota can metabolise glucosinolate to products other than isothiocyanates. Thus, for example *Citrobacter* Wye 1 could potentially have a negative affect as it does not produce high levels of ITCs as they undergo a detoxification to form an unknown metabolite [29]. It is likely that the bacterial metabolism of glucosinolates is diverse and it is important to identify degraders that produce ITCs as well those that produce other products in order to gain a picture of what is beneficial in relation to both human gut metabolism and agriculture.

5. Conclusion

Studying the diversity of bacteria that can metabolise glucosinolates is required in order to understand how glucosinolates can benefit both human health and agriculture. In this respect it is of importance to expand our knowledge of the bacterial myrosinase enzymes. So far cmyr is the only characterized bacterial myrosinase in terms of both its gene and protein.

In summary, this study investigated the feasibility of generating large amounts of recombinant cmyr together with some characterisation. Future structural studies will give an insight into the nature of this novel GH3 class of bacterial myrosinases.

Supplementary Materials: The following are available online, Figure S1: The multiple alignment of cmyr with some of the identified myrosinases. The abbreviations were used for myrosinase from: CMYR; *Citrobacter* Wye1,

BREBR (Q95X01); *B. brassicae* (aphid), ARATH (P37702); *A. thaliana*, EUTJA (Q4AE75); *E. japonicum*, SINAL (P29092); *S. alba*, BRANA (Q00326); *B. napus*, BRAO (A6XG32); *B. oleracea var. alboglabra*. The default settings were used for Clustal Omega. The entry numbers of sequences at Uniprot database are given in brackets. SDW motif was highlighted. The proton donor and nucleophile in *S. alba* myrosinase are in Gln 207 (Q) and Glu 426 (E) positions, respectively and they are highlighted as well. Figure S2: The MS spectrum for cmyr. Appropriate fragments of cmyr were determined in positive MS-TOF.

Funding: Quadram Institute Bioscience is supported by the Biotechnological and Biological Sciences Research Council via Institute Strategic grants BB/J004529/1, BB/J004545/1, BB/R012490/1 and BB/R012512/1. Fatma Cebeci was supported by a scholarship funded by the Republic of Turkey Ministry of National Education.

Conflicts of Interest: None.

Availability of Data: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Author' Contributions: Conceptualization, J.T.R., R.M. and A.N.; methodology, F.C. and M.M; validation, F.C.; formal analysis, F.C.; investigation, F.C.; resources, A.N.; writing—original draft preparation, F.C.; writing—review and editing, J.T.R., M.M, R.M. and A.N.; visualization, F.C and J.T.R.; supervision, J.T.R, M.M, R.M. and A.N.; project administration, A.N.; funding acquisition, A.N. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: We wish to thank to Dr. Gerhard Saalbach (John Innes Centre, Norwich) for his support to perform LCMS analysis.

References

[1] A. M. Bones and J. T. Rossiter, "The enzymic and chemically induced decomposition of glucosinolates," *Phytochemistry*. 2006, doi: 10.1016/j.phytochem.2006.02.024.

[2] R. Kissen, J. T. Rossiter, and A. M. Bones, "The 'mustard oil bomb': Not so easy to assemble?! Localization, expression and distribution of the components of the myrosinase enzyme system," *Phytochemistry Reviews*, vol. 8, no. 1. pp. 69–86, 2009, doi: 10.1007/s11101-008-9109-1.

[3] R. Mithen and E. Ho, "Isothiocyanates for Human Health," *Molecular Nutrition and Food Research*, vol. 62, no. 18. 2018, doi: 10.1002/mnfr.201870079.

[4] S. Novío, M. E. Cartea, P. Soengas, M. Freire-Garabal, and M. J. Núñez-Iglesias, "Effects of Brassicaceae isothiocyanates on prostate cancer," *Molecules*, vol. 21, no. 5. 2016, doi: 10.3390/molecules21050626.

[5] F. S. Chew, "Searching for Defensive Chemistry in the Cruciferae, or, Do Glucosinolates Always Control Interactions of Cruciferae with Their Potential Herbivores and Symbionts? No!," in *Chemical Mediation of Coevolution*, 1988, pp. 81–112.

[6] F. S. Chew, "Biological Effects of Glucosinolates," 1988, pp. 155–181.

[7] L. J. Rivera-Vega, S. Krosse, R. M. de Graaf, J. Garvi, R. D. Garvi-Bode, and N. M. van Dam, "Allelopathic effects of glucosinolate breakdown products in Hanza [Boscia senegalensis (Pers.) Lam.] processing waste water," *Front. Plant Sci.*, vol. 6, no. JULY, 2015, doi: 10.3389/fpls.2015.00532.

[8] C. Neubauer, B. Heitmann, and C. Müller, "Biofumigation potential of Brassicaceae cultivars to Verticillium dahliae," *Eur. J. Plant Pathol.*, vol. 140, no. 2, pp. 341–352, 2014, doi: 10.1007/s10658-014-0467-9.

[9] B. M. Ngala, P. P. Haydock, S. Woods, and M. A. Back, "Biofumigation with Brassica juncea, Raphanus sativus and Eruca sativa for the management of field populations of the potato cyst nematode Globodera pallida," *Pest Manag. Sci.*, vol. 71, no. 5, pp. 759–769, 2015, doi: 10.1002/ps.3849.

[10] T. Sotelo, M. Lema, P. Soengas, M. E. Cartea, and P. Velasco, "In vitro activity of Glucosinolates and their degradation products against Brassica-pathogenic bacteria and fungi," *Appl. Environ. Microbiol.*, vol. 81, no. 1, pp. 432–440, 2015, doi: 10.1128/AEM.03142-14.

[11] U. Wittstock and M. Burow, "Tipping the scales - Specifier proteins in glucosinolate hydrolysis," *IUBMB Life*, vol. 59, no. 12. pp. 744–751, 2007, doi: 10.1080/15216540701736277.

[12] F. Geng *et al.*, "Allyl isothiocyanate arrests cancer cells in mitosis, and mitotic arrest in turn leads to apoptosis via Bcl-2 protein phosphorylation," *J. Biol. Chem.*, vol. 286, no. 37, pp. 32259–32267, 2011, doi: 10.1074/jbc.M111.278127.

[13] V. Yoxall, P. Kentish, N. Coldham, N. Kuhnert, M. J. Sauer, and C. Ioannides, "Modulation of hepatic cytochromes P450 and phase II enzymes by dietary doses of sulforaphane in rats: Implications for its chemopreventive activity," *Int. J. Cancer*, vol. 117, no. 3, pp. 356–362, 2005, doi: 10.1002/ijc.21191.

[14] W. P. Burmeister, S. Cottaz, H. Driguez, R. Iori, S. Palmieri, and B. Henrissat, "The crystal structures of Sinapis alba myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase," *Structure*, 1997, doi: 10.1016/s0969-2126(97)00221-9.

[15] W. P. Burmeister, S. Cottaz, P. Rollin, A. Vasella, and B. Henrissat, "High resolution x-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base," *J. Biol. Chem.*, vol. 275, no. 50, pp. 39385–39393, 2000, doi: 10.1074/jbc.M006796200.

[16] B. Pontoppidan, B. Ekbom, S. Eriksson, and J. Meijer, "Purification and characterization of myrosinase from the cabbage aphid (Brevicoryne brassicae), a brassica herbivore," *Eur. J. Biochem.*, vol. 268, no. 4, pp. 1041–1048, 2001, doi: 10.1046/j.1432-1327.2001.01971.x.

[17] A. M. E. Jones, M. Bridges, A. M. Bones, R. Cole, and J. T. Rossiter, "Purification and characterisation of a non-plant myrosinase from the cabbage aphid Brevicoryne brassicae (L.)," *Insect Biochem. Mol. Biol.*, vol. 31, no. 1, pp. 1–5, 2001, doi: 10.1016/S0965-1748(00)00157-0.

[18] A. M. E. Jones, P. Winge, A. M. Bones, R. Cole, and J. T. Rossiter, "Characterization and evolution of a myrosinase from the cabbage aphid Brevicoryne brassicae," *Insect Biochem. Mol. Biol.*, vol. 32, no. 3, pp. 275–284, 2002, doi: 10.1016/S0965-1748(01)00088-1.

[19] F. Francis, G. Lognay, J. P. Wathelet, and E. Haubruge, "Characterisation of aphid myrosinase and degradation studies of glucosinolates," *Arch. Insect Biochem. Physiol.*, vol. 50, no. 4, pp. 173–182, 2002, doi: 10.1002/arch.10042.

[20] H. Husebye *et al.*, "Crystal structure at 1.1 Å resolution of an insect myrosinase from Brevicoryne brassicae shows its close relationship to β-glucosidases," *Insect Biochem. Mol. Biol.*, vol. 35, no. 12, pp. 1311–1320, 2005, doi: 10.1016/j.ibmb.2005.07.004.

[21] F. Beran *et al.*, "Phyllotreta striolata flea beetles use host plant defense compounds to create their own glucosinolate-myrosinase system," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 20, pp. 7349–7354, 2014, doi: 10.1073/pnas.1321781111.

[22] N. Tani, M. Ohtsuru, and T. Hata, "Purification and General Characteristics of Bacterial Myrosinase Produced by Enterobacter cloacae," *Agric. Biol. Chem.*, vol. 38, no. 9, pp. 1623–1630, 1974, doi: 10.1271/bbb1961.38.1623.

[23] M. Ll. Palop, J. P. Smiths, and B. ten Brink, "Degradation of sinigrin by Lactobacillus agilis strain R16," *Int. J. Food Microbiol.*, vol. 26, no. 2, pp. 219–229, 1995, doi: 10.1016/0168-1605(95)00123-2.

[24] G. H. Meulenbeld and S. Hartmans, "Thioglucosidase activity from Sphingo bacterium sp. strain OTG1," *Appl. Microbiol. Biotechnol.*, vol. 56, no. 5–6, pp. 700–706, 2001, doi: 10.1007/s002530100726.

[25] V. Luang-In, A. Narbad, C. Nueno-Palop, R. Mithen, M. Bennett, and J. T. Rossiter, "The metabolism of methylsulfinylalkyl- and methylthioalkyl-glucosinolates by a selection of human gut bacteria," *Mol. Nutr. Food Res.*, vol. 58, no. 4, pp. 875–883, 2014, doi: 10.1002/mnfr.201300377.

[26] M. Ohtsuru, I. Tsuruo, and T. Hata, "Studies on Fungous Myrosinase," *Agric. Biol. Chem.*, vol. 33, no. 9, pp. 1309–1325, 1969, doi: 10.1080/00021369.1969.10859462.

[27] J. P. Smits, W. Knol, and J. Bol, "Glucosinolate degradation by Aspergillus clavatus and Fusarium oxysporum in liquid and solid-state fermentation," *Appl. Microbiol. Biotechnol.*, vol. 38, no. 5, pp. 696–701, 1993, doi: 10.1007/BF00182812.

[28] N. Rakariyatham, B. Butr-Indr, H. Niamsup, and L. Shank, "Improvement of myrosinase activity of Aspergillus sp. NR4617 by chemical mutagenesis," *Electron. J. Biotechnol.*, vol. 9, no. 4, pp. 341–347, 2006, doi: 10.2225/vol9-issue4-fulltext-13.

[29] A. Narbad and J. T. Rossiter, "Gut Glucosinolate Metabolism and Isothiocyanate Production," *Molecular Nutrition and Food Research*, vol. 62, no. 18. 2018, doi: 10.1002/mnfr.201700991.

[30] J. A. Mullaney, J. Ansell, W. J. Kelly, and J. A. Heyes, "Biotransformation of glucosinolates from a bacterial perspective," *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, vol. 8. 2013, doi: 10.1079/PAVSNNR20138034.

[31] G. Rouzaud, S. Rabot, B. Ratcliffe, and A. J. Duncan, "Influence of plant and bacterial myrosinase activity on the metabolic fate of glucosinolates in gnotobiotic rats," *Br. J. Nutr.*, 2003, doi: 10.1079/bjn2003900.

[32] L. Elfoul, S. Rabot, N. Khelifa, A. Quinsac, A. Duguay, and A. Rimbault, "Formation of allyl isothiocyanate from sinigrin in the digestive tract of rats monoassociated with a human colonic strain of Bacteroides thetaiotaomicron," *FEMS Microbiol. Lett.*, vol. 197, no. 1, pp. 99–103, 2001, doi: 10.1016/S0378-1097(01)00093-3.

[33] A. Albaser *et al.*, "Discovery of a Bacterial Glycoside Hydrolase Family 3 (GH3) β-Glucosidase with Myrosinase Activity from a Citrobacter Strain Isolated from Soil," *J. Agric. Food Chem.*, vol. 64, no. 7, pp. 1520–1527, 2016, doi: 10.1021/acs.jafc.5b05381.

[34] C. S. Liou *et al.*, "A Metabolic Pathway for Activation of Dietary Glucosinolates by a Human Gut Symbiont," *Cell*, vol. 180, no. 4, pp. 717-728.e19, 2020, doi: 10.1016/j.cell.2020.01.023.
[35] A. M. Bones and J. T. Rossiter, "The myrosinase-glucosinolate system, its organisation and biochemistry," *Physiologia Plantarum*, vol. 97, no. 1. pp. 194–208, 1996, doi: 10.1111/j.1399-3054.1996.tb00497.x.
[36] R. Bhat and D. Vyas, "Myrosinase: insights on structural, catalytic, regulatory, and environmental interactions," *Critical Reviews in Biotechnology*, vol. 39, no. 4. pp. 508–523, 2019, doi: 10.1080/07388551.2019.1576024.
[37] F. Sievers *et al.*, "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega," *Mol. Syst. Biol.*, vol. 7, 2011, doi: 10.1038/msb.2011.75.
[38] V. Luang-In, "Influence of Human Gut Microbiota on the Metabolic Fate of Glucosinolates," Imperial College London, 2013.
[39] R. Bhat, T. Kaur, M. Khajuria, R. Vyas, and D. Vyas, "Purification and Characterization of a Novel Redox-Regulated Isoform of Myrosinase (β-Thioglucoside Glucohydrolase) from Lepidium latifolium L.," *J. Agric. Food Chem.*, vol. 63, no. 47, pp. 10218–10226, 2015, doi: 10.1021/acs.jafc.5b04468.
[40] S. Chen and B. A. Halkier, "Functional expression and characterization of the myrosinase MYR1 from Brassica napus in Saccharomyces cerevisiae," *Protein Expr. Purif.*, vol. 17, no. 3, pp. 414–420, 1999, doi: 10.1006/prep.1999.1158.
[41] M. Bridges *et al.*, "Spatial organization of the glucosinolate-myrosinase system in brassica specialist aphids is similar to that of the host plant," *Proc. R. Soc. B Biol. Sci.*, vol. 269, no. 1487, 2002, doi: 10.1098/rspb.2001.1861.
[42] E. Kazana *et al.*, "The cabbage aphid: A walking mustard oil bomb," *Proc. R. Soc. B Biol. Sci.*, vol. 274, no. 1623, 2007, doi: 10.1098/rspb.2007.0237.
[43] X. Li and M. M. Kushad, "Purification and characterization of myrosinase from horseradish (Armoracia rusticana) roots," *Plant Physiol. Biochem.*, vol. 43, no. 6, pp. 503–511, 2005, doi: 10.1016/j.plaphy.2005.03.015.
[44] "UK Soil Observatory." http://mapapps2.bgs.ac.uk/ukso/home.html (accessed Oct. 15, 2021).
[45] D. C. James and J. T. Rossiter, "Development and characteristics of myrosinase in Brassica napus during early seedling growth," *Physiol. Plant.*, vol. 82, no. 2, pp. 163–170, 1991, doi: 10.1111/j.1399-3054.1991.tb00076.x.
[46] R. Bernardi, M. G. Finiguerra, A. A. Rossi, and S. Palmieri, "Isolation and biochemical characterization of a basic myrosinase from ripe Crambe abyssinica seeds, highly specific for epi-progoitrin," *J. Agric. Food Chem.*, vol. 51, no. 9, pp. 2737–2744, 2003, doi: 10.1021/jf020796g.
[47] M. Sabaty *et al.*, "Detrimental effect of the 6 His C-terminal tag on YedY enzymatic activity and influence of the TAT signal sequence on YedY synthesis," *BMC Biochem.*, vol. 14, no. 1, 2013, doi: 10.1186/1471-2091-14-28.
[48] Y. Aslantas and N. B. Surmeli, "Effects of N-terminal and C-terminal polyhistidine tag on the stability and function of the thermophilic P450 CYP119," *Bioinorg. Chem. Appl.*, vol. 2019, 2019, doi: 10.1155/2019/8080697.
[49] Y. L. Lee, M. S. Su, T. H. Huang, and J. F. Shaw, "C-terminal His-tagging results in substrate specificity changes of the thioesterase I from Escherichia coli," *JAOCS, J. Am. Oil Chem. Soc.*, vol. 76, no. 10, pp. 1113–1118, 1999, doi: 10.1007/s11746-999-0082-7.
[50] F. Cebeci, "The metabolism of plant glucosinolates by gut bacteria," University of East Anglia, 2017.

Figure Captions

Fig. 1. Glucosinolate degradation by myrosinase [29].

Fig. 2. SDS-PAGE and Western Blot of soluble and insoluble extracts from *E. coli* BL21(DE3) expressing pET28b-*cmyr*. CFEs were run on 4-12% Bis-Tris gel under denaturing, reducing conditions in MES buffer.10 µg of protein was loaded per lane. M: See Blue Protein Marker (Invitrogen). 1-2; soluble and insoluble CFEs of induced (25°C, 4 h) cells respectively. 3-4; Western blot (using anti-His tag antibody) within the same order of the protein gel.

Fig. 3. SDS-PAGE of the Ni-NTA purified cmyr protein from *E. coli* BL21 (DE3) expressing pET28b-*cmyr*. Proteins were run on 4-12% Bis-Tris Gel under denaturing, reducing conditions in MES buffer. 3.25 µl of fractions were loaded per lane. L; lysate, FT; flow through, W; wash, E; eluate, M; See Blue protein marker (Invitrogen). The expected protein size of cmyr with N- and C-terminal 6X-His tags was 75 kDa.

Fig. 4. pH optimum range of recombinant cmyr myrosinase activity in 20 mM citrate phosphate buffer. The myrosinase activity was determined at 37°C after 30 min incubation measuring glucose release by God-Perid Assay. Error bars = standard deviation (n=3).

Fig. 5. Temperature optimisation of cmyr in 20 mM citrate phosphate buffer pH 6.0. The myrosinase activity was determined after 30 min incubation measuring glucose release by God-Perid Assay. Error bars = standard deviation (n=3).

Fig. 6. Kinetic analysis of cmyr using sinigrin as substrate. Specific activity (mmol/min/mg protein) of cmyr (24 ng incubated) at different concentrations of sinigrin (0.1-5 mM) was used to prepare non-linear Michaelis Menten plot.

Fig. 7. Half life of cmyr activity. The myrosinase activity was determined at 25°C in 20 mM citrate phosphate buffer measuring glucose release by God-Perid assay (SD,n=3).

Fig. 8. Potential mechanism of cmyr catalyzed glucosinolate hydrolysis.

Tables and Their Captions

Table 1. Primers and conditions used in the PCR experiments to amplify the cmyr.

Name	Primer Sequence (5`-3`)	T_A, time, cycles	T_E, time
CMYR-F	GGAACATATGCTCACTGCTTTTAAGA	54°C, 30 s, 5X	72°C, 1 min
CMYR-R	CAGA**CTCG**AG**A**CGTGTCAGTCCGAAT	68°C, 30 s, 20X	

The restriction sites are NdeI and XhoI for CMYR-F and CMYR-R primers respectively and these sites are underlined. The altered nucleotides to introduce the restriction sites are given in bold. The inserted nucleotide is highlighted. T_A; annealing temperature, cycles; cycles used in annealing step of PCR, T_E; extension temperature and time used in PCR. The annealing step was performed using Tm of 100% sequence match part of primers first (5X) and then using the Tm of entire primer (20X).

Table 2. The myrosinase activity of cmyr towards different glucosinolates

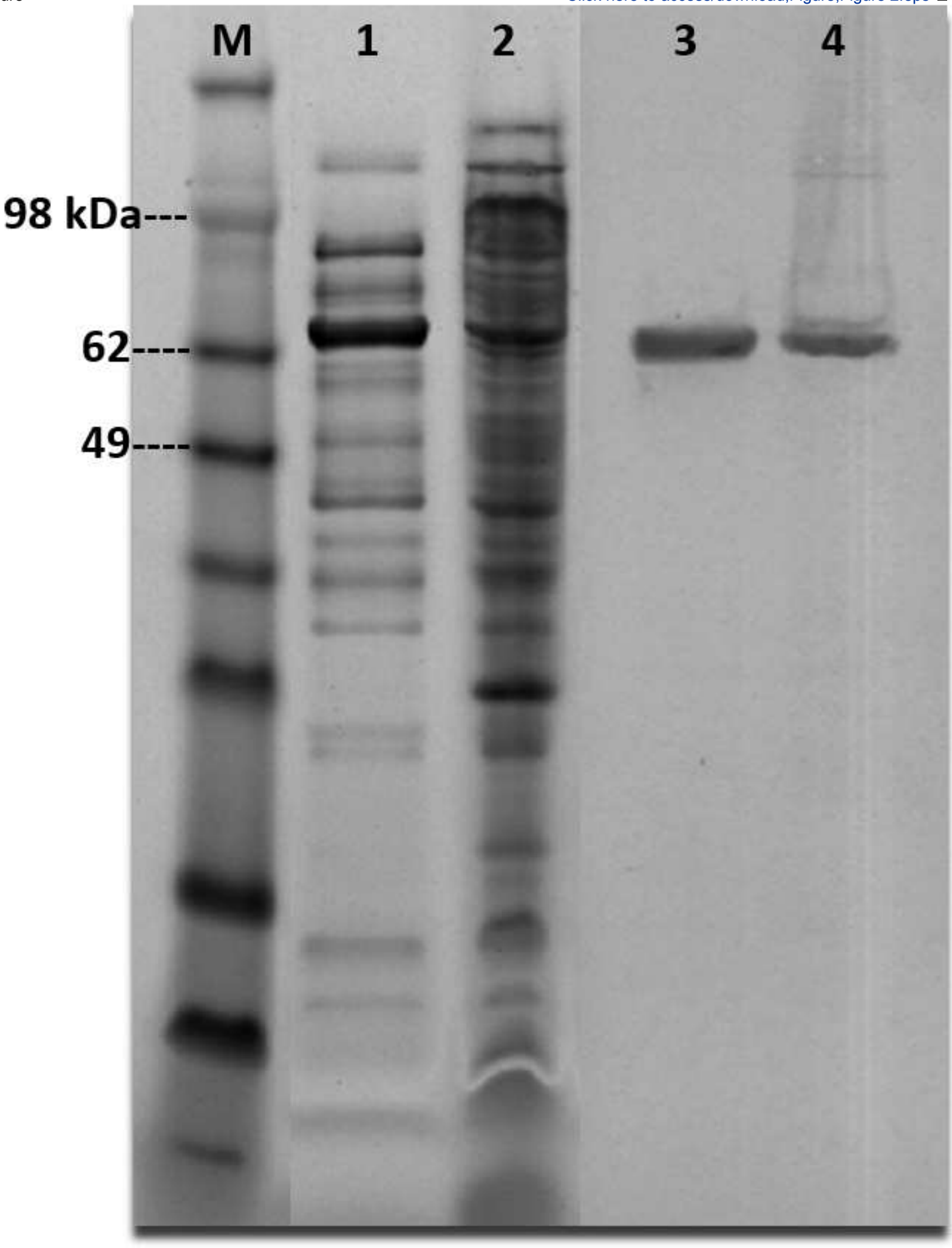
Glucosinolates	Specific activity (μmol/min/mg protein)
Sinigrin	47.31 ± 1.25
Glucoiberin	39.96 ± 3.90
Progoitrin	35.42 ± 1.72
Glucoerucin	30.48 ± 4.60
Glucoraphanin	5.28 ± 0.38
Glucotropaeolin	2.38 ± 0.10
Gluconasturtiin	ND
Glucobrassicin	ND

Myrosinase activity was determined at 25°C in 20 mM citrate phosphate buffer and measuring glucose release by the God-Perid assay. All tests were performed using the same batch of purified enzyme on same day (SD, n=3). ND, not detected.

Figure

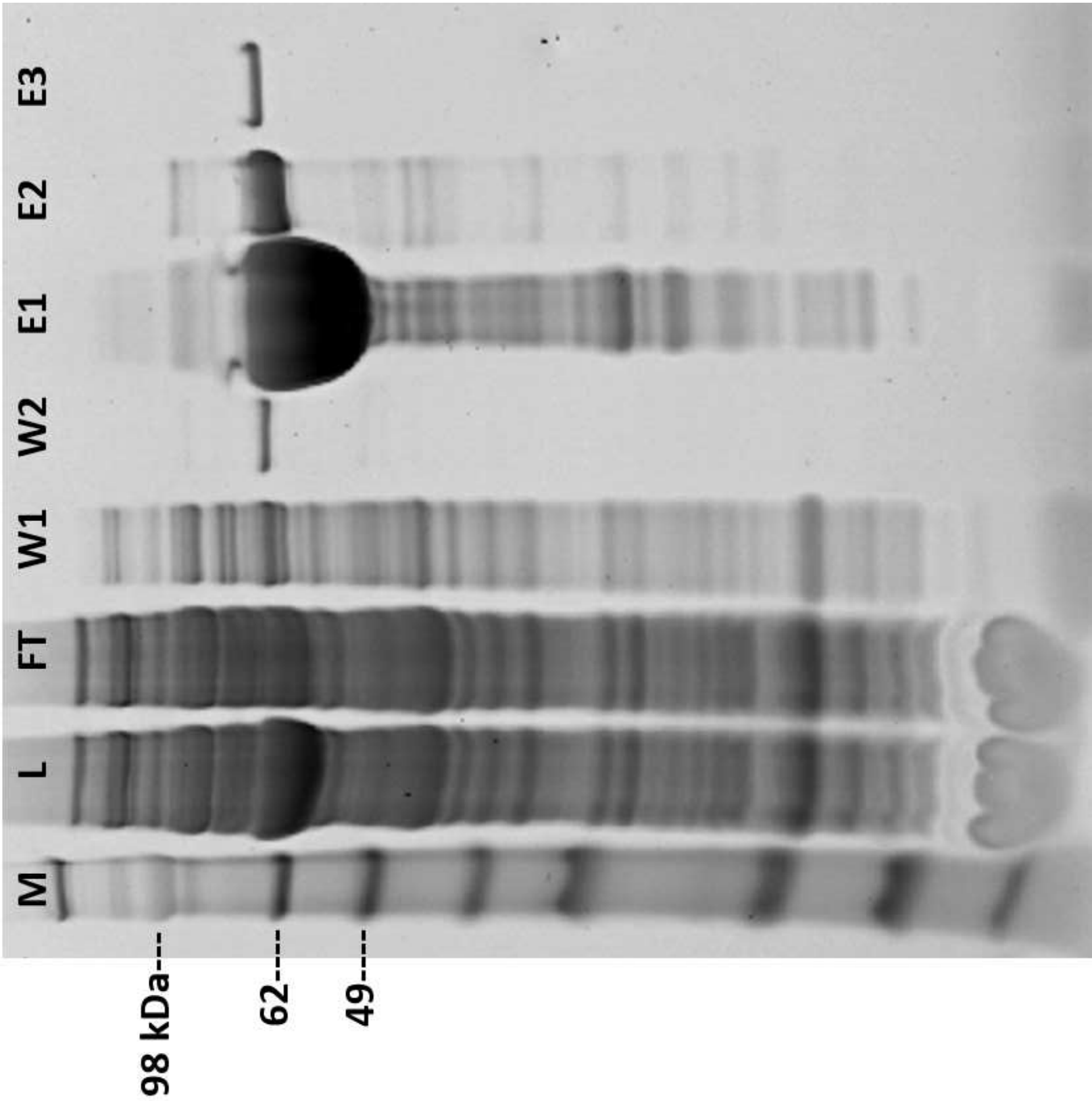
Figure

Click here to access/download;Figure;Figure 2.eps



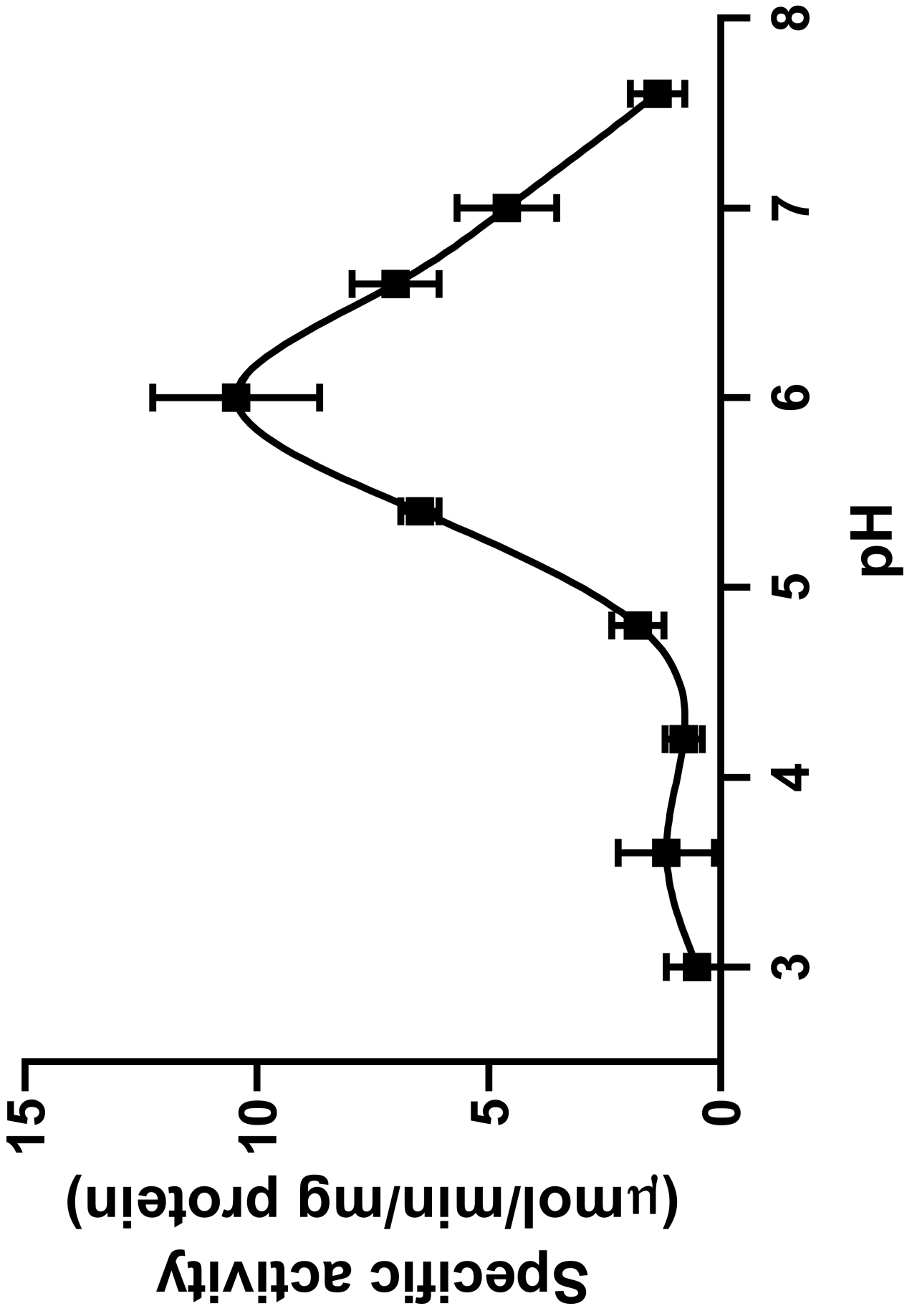
Figure

Click here to access/download;Figure;Figure 3.eps



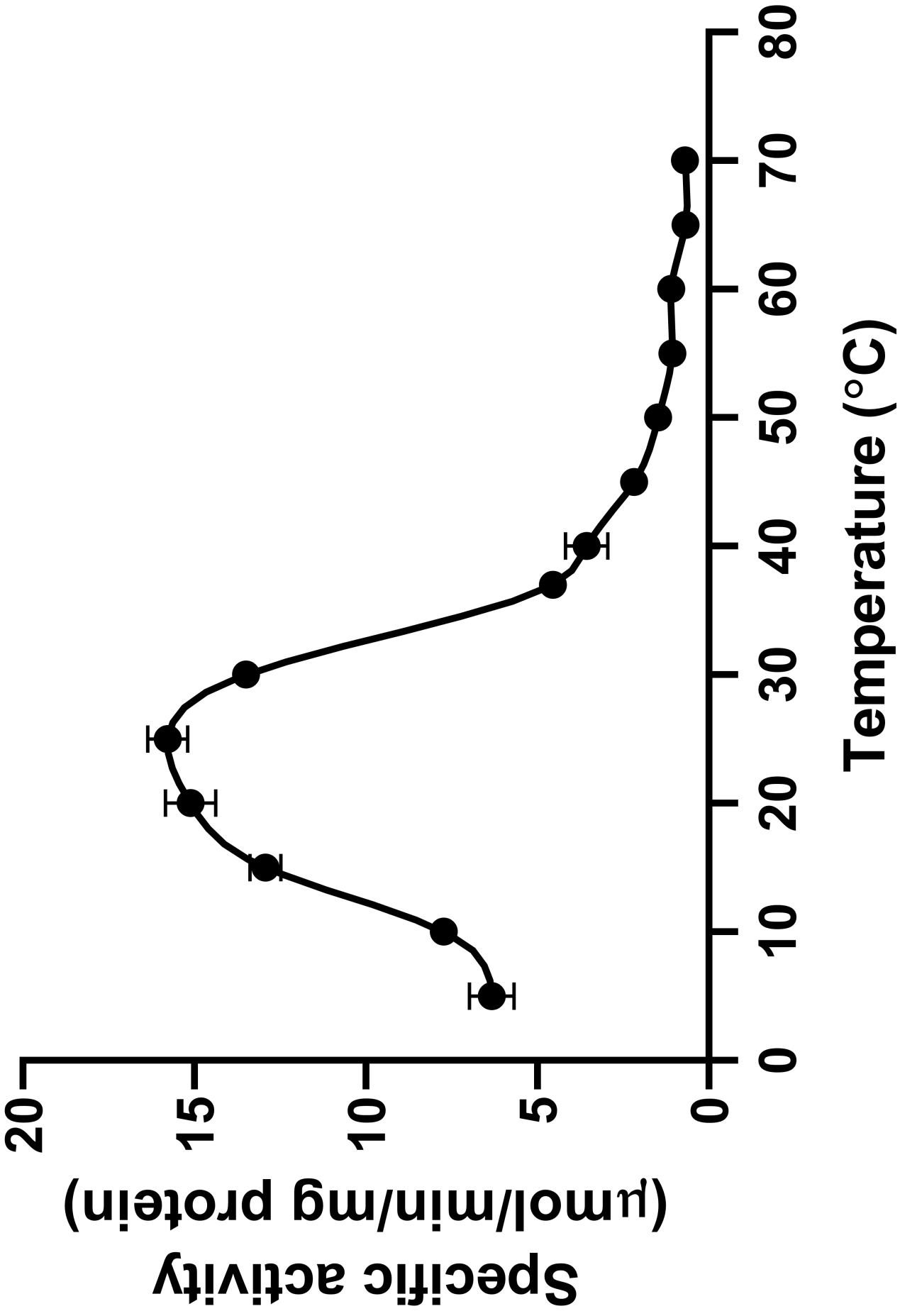
Figure

Click here to access/download;Figure;Figure 4.eps



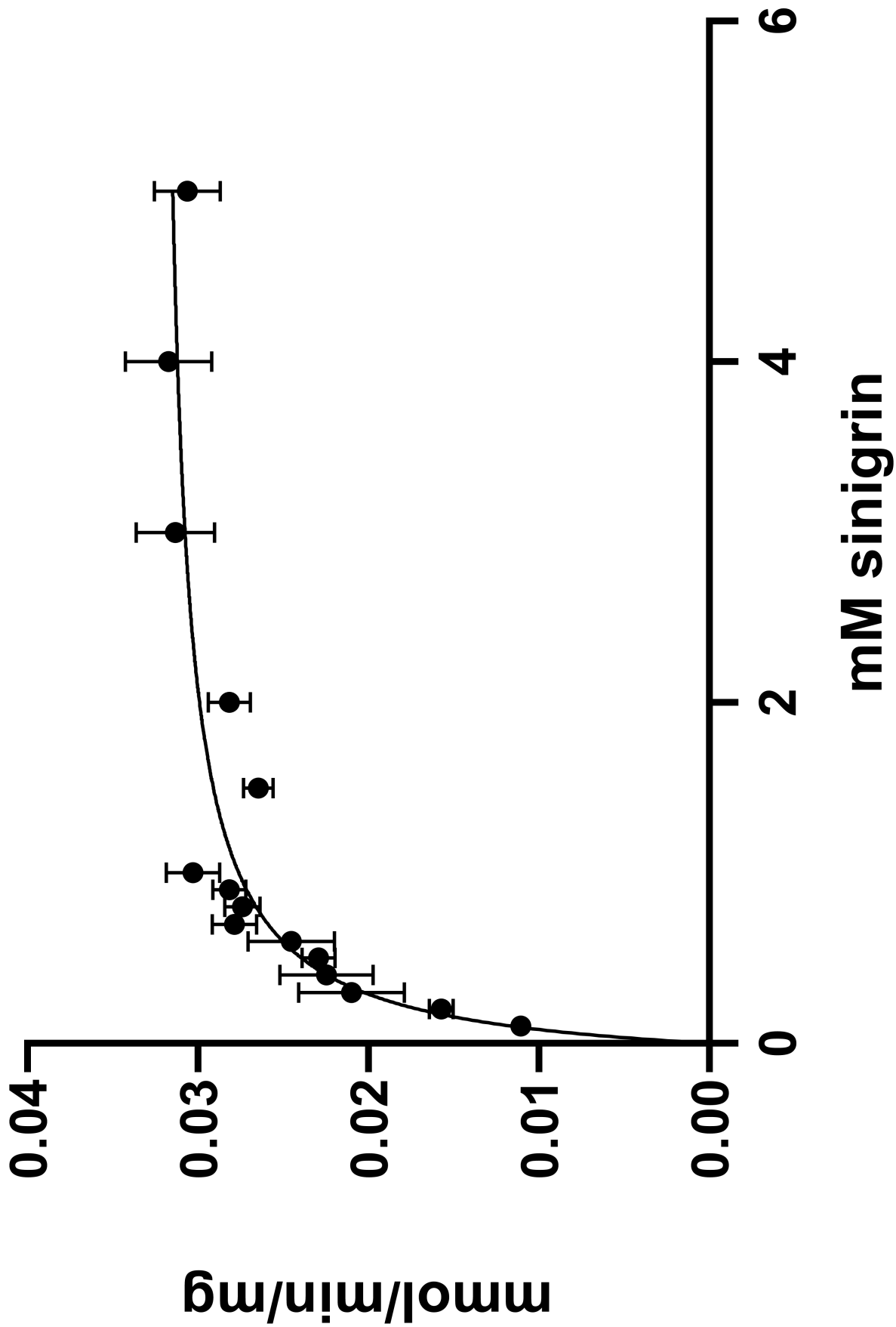
Figure

Click here to access/download;Figure;Figure 5.eps



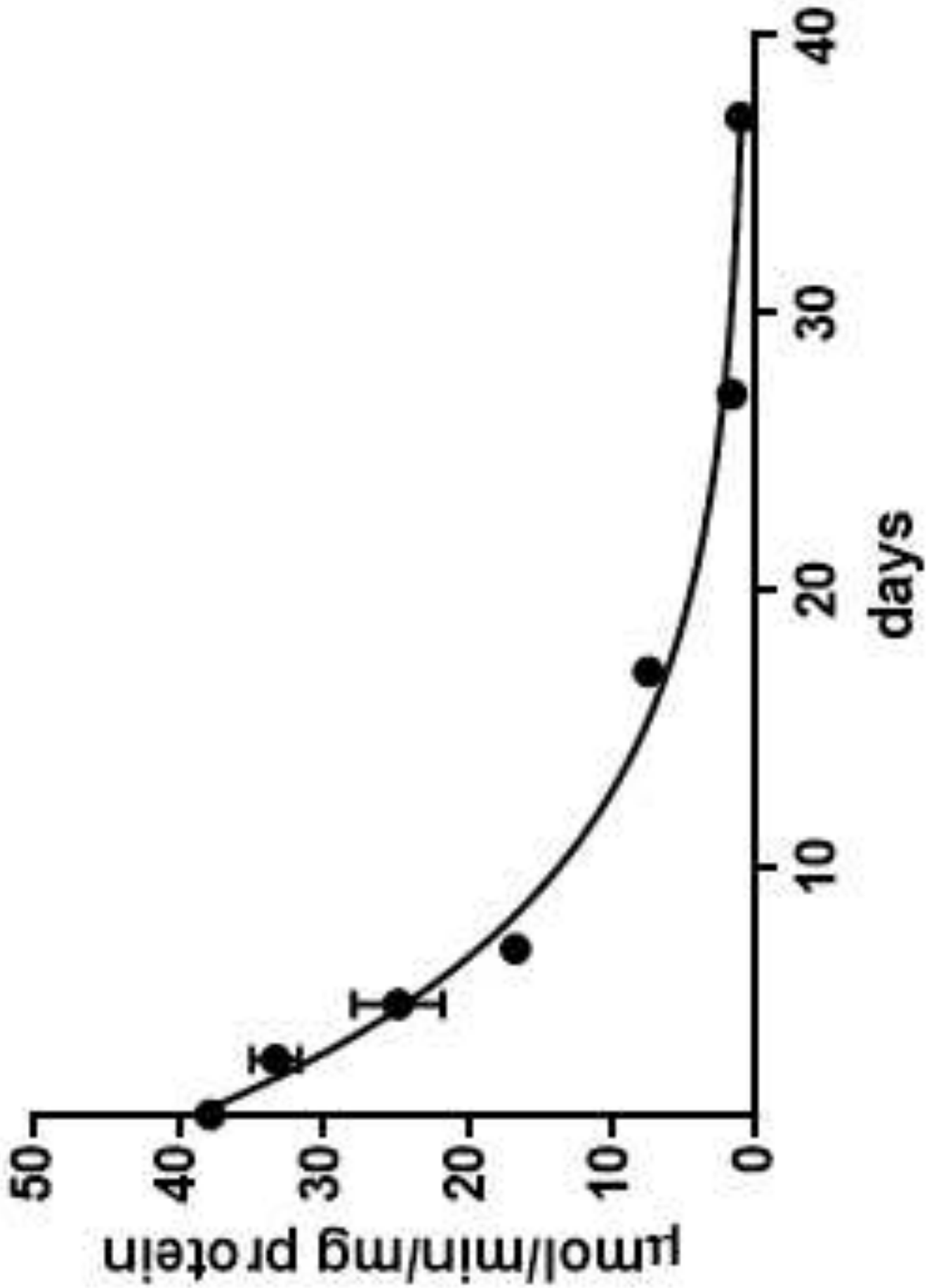
Figure

Click here to access/download;Figure;Figure 6.eps



Figure

Click here to access/download;Figure;Figure 7.eps



Figure

Click here to access/download;Figure;Figure 8.eps