Parasites Hosts Dis 2024;62(3):330-341
https//doi.org/10.3347/PHD.24026
pISSN 2982-5164 • eISSN 2982-6799

Original Article

Identification of the protease inhibitory domain of *Trichinella spiralis* novel cystatin (TsCstN)

Thassanee Yuthithum[1], Orawan Phuphisut[1], Onrapak Reamtong[2], Nathamon Kosoltanapiwat[3], Salisa Chaimon[4], Porntida Kobpornchai[5], Charin Thawornkuno[2], Preeyarat Malaithong[1], Orathai Sawatdichaikul[6,*], Poom Adisakwattana[1,*]

[1]Department of Helminthology, Faculty of Tropical Medicine, Mahidol University, Ratchathewi, Bangkok 10400, Thailand; [2]Department of Molecular Tropical Medicine and Genetics, Faculty of Tropical Medicine, Mahidol University, Ratchathewi, Bangkok 10400, Thailand; [3]Department of Microbiology and Immunology, Faculty of Tropical Medicine, Mahidol University, Ratchathewi, Bangkok 10400, Thailand; [4]Department of Preclinical Science, Faculty of Medicine, Thammasat University, Klongluang, Pathumthani 12120, Thailand; [5]Department of Parasitology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkoknoi, Bangkok 10700, Thailand; [6]Institute of Food Research and Product Development, Kasetsart University, Chatuchak, Bangkok 10900, Thailand

Received: 28 March 2024
Accepted: 3 July 2024

***Correspondence**
(os, orathai.saw@ku.th;
pa, poom.adi@mahidol.ac.th)

Citation
Yuthithum T, Phuphisut O, Reamtong O, Kosoltanapiwat N, Chaimon S, Kobpornchai P, Thawornkuno C, Malaithong P, Sawatdichaikul O, Adisakwattana P. Identification of the protease inhibitory domain of *Trichinella spiralis* novel cystatin (TsCstN). Parasites Hosts Dis 2024;62(3):330-341.

Abstract

The *Trichinella spiralis* novel cystatin (TsCstN) inhibits cathepsin L (CatL) activity and inflammation of macrophages during lipopolysaccharide (LPS) induction. To identify the protease inhibitory region, this study applied an in silico modeling approach to simulate truncation sites of TsCstN (Ts01), which created four truncated forms, including TsCstN$^{\Delta 1\text{-}39}$ (Ts02), TsCstN$^{\Delta 1\text{-}71}$ (Ts03), TsCstN$^{\Delta 1\text{-}20,\ \Delta 73\text{-}117}$ (Ts04), and TsCstN$^{\Delta 1\text{-}20,\ \Delta 42\text{-}117}$ (Ts05). The superimposition of these truncates modeled with AlphaFold Colab indicated that their structures were more akin to Ts01 than those modeled with I-TASSER. Moreover, Ts04 exhibited the closest resemblance to the structure of Ts01. The recombinant Ts01 (rTs01) and truncated proteins (rTs02, rTs03, and rTs04) were successfully expressed in a prokaryotic expression system while Ts05 was synthesized, with sizes of approximately 14, 12, 8, 10, and 2.5 kDa, respectively. When determining the inhibition of CatL activity, both rTs01 and rTs04 effectively reduced CatL activity in vitro. Thus, the combination of the α1 and L1 regions may be sufficient to inhibit CatL. This study provides comprehensive insights into TsCstN, particularly regarding its protein function and inhibitory domains against CatL.

Keywords: *Trichinella spiralis*, novel cystatin, in silico analysis, protein truncation, protease inhibitor, cathepsin L

Introduction

Trichinella spiralis is a zoonotic nematode that causes trichinellosis (or trichinosis) in humans and can infect various mammals worldwide [1]. The parasite can be transmitted by consuming raw or undercooked infected meat, especially pork muscle that contains larvae (TsL1). After exposure to gastric juice in the stomach, TsL1 larvae are released from the muscle tissue and then penetrate the small intestinal wall and develop into adult worms. Adult males and females mate, after which the females release newborn larvae (NBL) that migrate via the blood or lymphatic circulation to muscles and particularly to highly oxygenated skeletal muscles [2]. To survive long-term in their hosts, *T. spiralis* larvae establish chronic infections in their host muscles by modulating and escaping potentially destructive immune responses. This is achieved by anatomical seclusion involving intracellular in-

Author contributions
Conceptualization: Phuphisut O, Reamtong O, Kosoltanapiwat N, Thawornkuno C, Sawatdichaikul O, Adisakwattana P
Data curation: Yuthithum T, Phuphisut O, Sawatdichaikul O, Adisakwattana P
Formal analysis: Yuthithum T, Phuphisut O, Chaimon S, Kobpornchai P, Sawatdichaikul O, Adisakwattana P
Funding acquisition: Adisakwattana P
Investigation: Yuthithum T, Phuphisut O, Kobpornchai P, Sawatdichaikul O, Adisakwattana P
Methodology: Yuthithum T, Phuphisut O, Thawornkuno C, Malaithong P, Sawatdichaikul O, Adisakwattana P
Project administration: Malaithong P, Adisakwattana P
Resources: Phuphisut O, Malaithong P, Adisakwattana P
Software: Yuthithum T, Phuphisut O, Chaimon S, Sawatdichaikul O, Adisakwattana P
Supervision: Phuphisut O, Reamtong O, Kosoltanapiwat N, Sawatdichaikul O, Adisakwattana P
Validation: Yuthithum T, Phuphisut O, Sawatdichaikul O, Adisakwattana P
Visualization: Yuthithum T, Phuphisut O, Sawatdichaikul O, Adisakwattana P
Writing – original draft: Yuthithum T, Phuphisut O, Sawatdichaikul O, Adisakwattana P
Writing – review & editing: Yuthithum T, Phuphisut O, Reamtong O, Kosoltanapiwat N, Thawornkuno C, Sawatdichaikul O, Adisakwattana P

Conflict of interest
The authors declare no conflict of interest related to this study.

ORCID
Thassanee Yuthithum (https://orcid.org/0009-0003-5449-1183)
Orathai Sawatdichaikul (https://orcid.org/0000-0001-9158-7924)
Poom Adisakwattana (https://orcid.org/0000-0002-0313-9900)

fection and nurse cell formation, as well as regulating host immune responses by releasing immunomodulatory molecules found in excretory–secretory products (ES) [2,3]. In some instances, TsL1-ES can modulate the immune response primarily through the TLR4-dependent pathway by decreasing the expression of this receptor [4]. The incubation of TsL1-ES with HEK cells transfected with mouse TLR4 (TLR4/MD2-CD14) can suppress LPS-mediated TLR4 activation, reducing proinflammatory cytokine production [5].

In our previous study, recombinant *T. spiralis* novel cystatin (rTsCstN) inhibited the activity of cathepsin L (CatL) in vitro and ex vivo in mouse bone marrow-derived macrophages (mBMDMs). Moreover, rTsCstN exhibited anti-inflammatory properties in mBMDMs by suppressing lipopolysaccharide (LPS)-mediated proinflammatory cytokines, including tumor necrosis factor (TNF)-α, interleukin (IL)-1β, and interferon (IFN)-γ and by interfering with the antigen presentation process through the suppression of MHC class II expression [6]. In addition, the interaction between TsCstN-mBMDMs impaired the Th1 response, affecting downstream T-cell priming by lowering IFN-α and IL-12 levels and interfering with STAT4 signaling [7]. To achieve an in-depth understanding of properties and functions of TsCstN, the protease inhibitory domain of the protein should be identified and characterized and establish if future modifications or truncation of this protein could increase its activity and specificity against inflammatory diseases.

In previous studies, cystatin-derived peptides have been designed and used for therapeutic purposes, including anti-inflammatory, antimicrobial, and anticancer effects. The administration of peptides derived from filarial cystatin to mice with dextran sulfate sodium–induced colitis significantly reduced gross and histological pathological alterations and inflammatory responses [8]. The N-terminal fragment of human cystatin C exhibited strong antibacterial action against several bacterial species, including methicillin-resistant *S. aureus*, frequently isolated from infected wounds [9]. A small cystatin C peptide also demonstrated anticancer effects by decreasing cell growth, proliferation, and migration and increasing apoptosis in melanoma cells in vitro [10].

In this study, we identified the protease inhibitory reactive regions of TsCstN by using an in silico truncation modeling approach. Full-length TsCstN was employed to simulate secondary (2D) and tertiary (3D) structures before determining the truncation locations based on the functional domains of the proteins. Subsequently, recombinant proteins of TsCstN and its truncates were expressed in a prokaryotic expression system or artificially synthesized, followed by an evaluation of their properties through the inhibition of human CatL. The results of this study will help in the future development of biotherapeutic medications for immune-mediated disorders and chronic inflammatory diseases.

Materials and Methods

Ethics statement

Not applicable.

Bioinformatic approach to design TsCstN protein truncates

In our previous study, TsCstN was shown to lack conserved cystatin sequences according to multiple protein alignments with its homologs [6]. Therefore, we gathered additional

cystatins derived from various helminths to reassess the multiple alignment and validate our earlier discovery [6]. Amino acid sequences of ten helminth cystatins were retrieved from the NCBI database (https://www.ncbi.nlm.nih.gov/genbank/); amino acid sequences and accession numbers are provided in Supplementary Table S1. Multiple sequence alignment was performed to compare the conserved sequence regions using ClustalX2 [11] and visualized using the Bio-Edit alignment editor [12].

Design of TsCstN truncates and protein structure simulation

The amino acid sequence of full-length TsCstN (Ts01) was used to determine the protein 2D structure predicted by five different programs: Jpred4 [13], PHD [14], Phyre2 [15], PSIPRED [12], and PredictProtein (https://predictprotein.org/). The TsCstN truncates, TsCstN$^{\Delta 1\text{-}39}$ (Ts02), TsCstN$^{\Delta 1\text{-}71}$ (Ts03), TsCstN$^{\Delta 1\text{-}20,\ \Delta 73\text{-}117}$ (Ts04), and TsCstN$^{\Delta 1\text{-}20,\ \Delta 42\text{-}117}$ (Ts05), were manually designed by removing protein domains based on the 2D structures and functional protein domains of Ts01 to generate Ts02–Ts05 were subsequently predicted as 2D structures using the same programs as above for Ts01 to confirm structure change and stability after truncation; predicted structures were then aligned using the Bio-Edit alignment editor [16]. The molecular weight and isoelectric point of Ts01–Ts05 were predicted using the ProtParam tool on the ExPASy server (http://www.expasy.org/tools/).

I-TASSER [17] and the AlphaFold2 method [18] were used to simulate the 3D structure of Ts01–Ts05. The surface area, volume, and 3D structures of these proteins were visualized using Biovia discovery studio (Discovery Studio Modeling Environment, Release 2022. Dassault Systèmes; San Diego, CA, USA: 2021), and the Ramachandran plot was evaluated to verify the quality of each structure using PROCHECK analysis [19]. Structures of truncated TsCstN were superimposed onto Ts01 using the PyMOL2 server (http://www.pymol.org/pymol). The overall conceptual framework of protein modeling is shown as a flow chart (Supplementary Fig. S1).

DNA amplification and cloning of truncated TsCstN DNAs

pET23b$^+$ carrying full-length TsCsN without a signal peptide (pET23b-Ts01) was prepared in our previous publication [6] and was used as the DNA template for amplifying truncated TsCstN (Ts02–Ts04) DNAs. The specific primers were designed to amplify the encoding DNA region of each truncate, and the *Bam*HI and *Hind*III restriction sites were incorporated into the 5°C-end of the forward (fw) and reverse (rv) primers, respectively, for cloning into pET23b$^+$. Specific primers for amplifying Ts01–Ts04 are listed in Supplementary Table S2. The Ts05 peptide was synthesized by GenScript (Piscataway, NJ, USA).

The truncated TsCstN was amplified via conventional PCR in a total volume of 20 μl containing 2 μl of pET23b-Ts01 plasmid (approximately 1 ng), 1× Dream *Taq* DNA polymerase buffer, 0.2 mm of each dNTP, 1 U of Dream *Taq* DNA polymerases, and 100 nM of Fw- and Rv-specific primers of each truncate. PCR conditions consisted of 94°C for 5 min, 35 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 2 min and a final extension step at 72°C for 5 min. Results were checked via 3% gel electrophoresis. The target amplicon was excised from the gel and purified using a GenepHlow Gel/PCR kit (Geneaid Biotech, New Taipei, Taiwan). The pET23b$^+$ plasmid (Novagen-EMD Biosciences, Darmstadt, Germany) and purified Ts02, Ts03, or Ts04 amplicons were digested with *Bam*HI and *Hind*III at 37°C for 4 h

before ligation with T4 DNA ligase (Thermo Fisher Scientific, Waltham, MA, USA). The ligation reaction was transformed into *Escherichia coli* strain JM109 via heat-shock. The following day, 5 to 10 single colonies from each plate of truncated TsCstN were selected to confirm and validate DNA recombination using colony PCR, restriction endonuclease digestion, and DNA sequencing.

Expression and purification of full-length and truncated TsCstN proteins

Recombinant truncated TsCstN, the plasmids pET23b^{+} carrying Ts02, Ts03, or Ts04 (pET 23b-Ts02, pET23b-Ts03, and pET23b-Ts04, respectively) were isolated from *E. coli* strain JM109, and transformed *E. coli* strain BL21 (DE3). BL21 cells carrying plasmid pET23b-Ts01, -Ts02, -Tso3, or -Ts04 were used to express and purify recombinant proteins as previously described [6]. Recombinant mouse dihydrofolate reductase (rmDHFR) was produced using the procedures outlined in our previous study and employed as an irrelevant protein control [6].

Protease inhibitory assay

The Ts01–Ts05 proteins were evaluated for their inhibition of human CatL activity (Sigma-Aldrich, St Louis, MO, USA). rTs01 (129 μM), rTs02 (392 μM), rTs03 (243 μM), rTs04 (350 μM), and rTs05 (19.6 mM) were diluted to 10, 20, or 40 μm and incubated with assay buffer (100 mM sodium acetate, pH 5.5, 100 mM NaCl, 1 mM DTT) at 37°C for 30 min. Subsequently, the reaction was added to Nunc F96 MicroWell Black plate, followed by the addition of 0.5 nM CatL into the well and incubation at 37°C for 30 min. Then, 100 μl of ta fluorogenic CatL substrate (10 mM of Z-Phe-Arg-AMC-HCl) (Chem-Impex International, Wood Dale, IL, USA) was added per reaction. The fluorescence intensity was monitored at 37°C for 30 min using a fluorometer (Tecan Infinite M Plex, Tecan Austria GmbH Untersbergstr. 1A, A-5082 Grödig, Austria) with excitation and emission at 360 and 480 nm, respectively. The percentage of residual enzyme activity was determined. rmDHFR and E64 were used as irrelevant and positive inhibitor controls, respectively. Experiments were performed in duplicate with two independent experiments.

Statistical analysis

Statistical analyses were performed using GraphPad Prism 7 software (GraphPad, La Jolla, CA, USA). The significance between each group was analyzed using one-way ANOVA followed by multiple comparison tests. Individual data in each group are presented as the mean ± SD; a P-value < 0.05 was considered significant.

Results

Designing TsCstN truncates using a nonhomology structure strategy

A multiple sequence alignment of TsCstN (Ts01) with related cystatin homologs was performed using ClustalX2. The Ts01 amino acid sequence had a low conservation level. Compared with cystatin homologs, Ts01 did not contain any of the cystatin signature motif present in cystatin homologs (Supplementary Fig. S2). The structure template for the 3D protein simulation of Ts01 was obtained from the SWISS-MODEL server (https://swiss-

model.expasy.org/interactive PDB) [20], although the PDB database currently lacks a structural homolog for Ts01. Therefore, it was not possible to simulate Ts01 or create protein truncates using comparative homology modeling.

To design the truncate TsCstN proteins, the Ts01 2D structures were predicted with pred4, PHD, Phyre2, and PSIPRED. The Ts01 2D structure and functional protein domain were used to predict the sites for truncation locations that would produce TsCstN truncates. Four distinct protein truncates, Ts02, Ts03, Ts04, and Ts05, were generated in this study. Following a comparison with the Ts01 structure, the 2D structures of the truncates were predicted in the same manner as for Ts01. The 2D structures of the truncates aligned with Ts01 from all from five different algorithms (Supplementary Fig. S3). The amino acid sequences, 2D structure composition of Ts01, truncates, molecular weights (kDa), and isoelectric point are summarized in Supplementary Table S3.

3D structural simulation of TsCstN truncates to evaluate conformational alteration

I-TASSER and Alpha-Fold2 were used to model Ts01 3D structures and truncates, with both techniques predicting all structures comparably. In detail, Ts01 contained a full α1 helix (α1), loop1 (L1; β1β2), α2 helix (α2), an irregular coil, and Loop2 (L2; β3β4) (1–117 aa); Ts02 contained L1, α2, an irregular coil, and L2 without α1 (40–117 aa); Ts03 contained α2, an irregular coil, and L2 without α1 or L1 (72–117 aa); Ts04 contained α1 and L1 (21–72 aa); and Ts05 refers to α1 structure only (21–41 aa) (Supplementary Figs. S4, S5). Results from the Ramachandran plots and the supporting information of protein structure modeling, such as surface area and volume, were also used for analysis. The structures of proteins showing a favored region pf > 70% and a disallowed region pf < 2% were selected to align against Ts01 to select the closely related structure that was similar to the predicted structure. Based on the Ramachandran plot, the findings indicated that I-TASSER exhibited a lower precision in predicting 3D protein structure. Most protein structures displayed a small percentage of favored regions (< 70%) and a significantly higher percentage of disallowed regions (> 2%). Conversely, AlphaFold2 demonstrated a higher accuracy in predict-

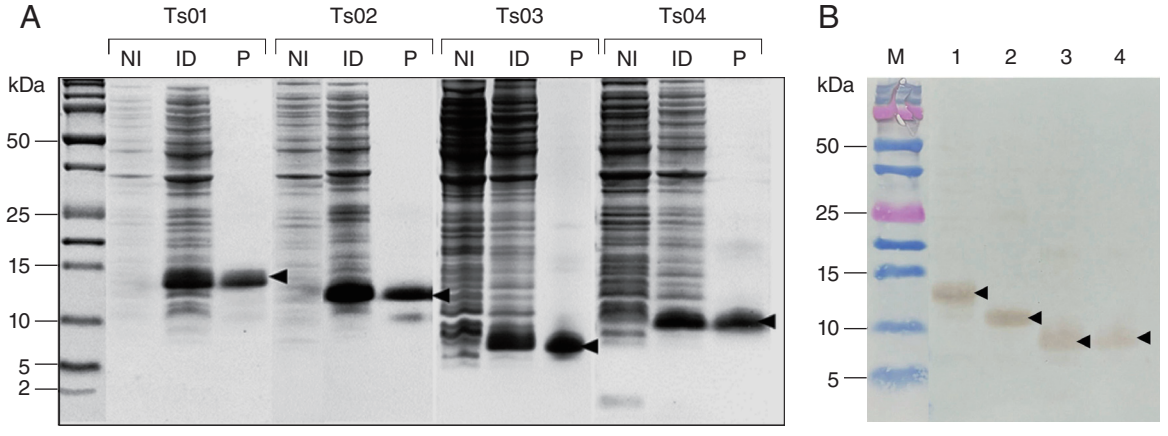

Fig. 1. The expression of rTsCstN (Ts01) and truncates Ts02, Ts03, and Ts04 after IPTG induction was analyzed via16.5% Tris-tricine SDS-PAGE (A), Western blot analysis (B) of rTs01 (lane 1), rTs02 (lane 2), rTs03 (lane 3), and rTs04 (lane 4) against mouse anti-His tag antibodies. M, Precision Plus Protein Dual Xtra standards (Bio-Rad).

ing 3D structures, featuring a notable percentage of the favored regions (>70%) and a minimal percentage of the disallowed regions (<2%). The 3D structure modeling of Ts04 demonstrated the best quality with the highest percentage of favored regions (Supplementary Tables S4, S5).

To validate the results of the 3D structure simulation, the truncated TsCstN (Ts02–Ts05) structures were superimposed on Ts01. The proteins modeled by AlphaFold Colab had a structure more comparable with that of Ts01 than those modeled with I-TASSER (Supplementary Figs. S6, S7). The structure of TS04 was the most comparable with that of Ts01 when modeled with AlphaFold2 (Supplementary Fig. S7C).

Production of recombinant proteins for functional analysis

Recombinant Ts01, -Ts02, -Ts03, and -Ts04 were overexpressed in *E. coli* in an insoluble form at molecular weights of approximately 14, 12, 8, and 10 kDa, respectively (Fig. 1A). Purified recombinant proteins were completely refolded and solubilized in 1× PBS before

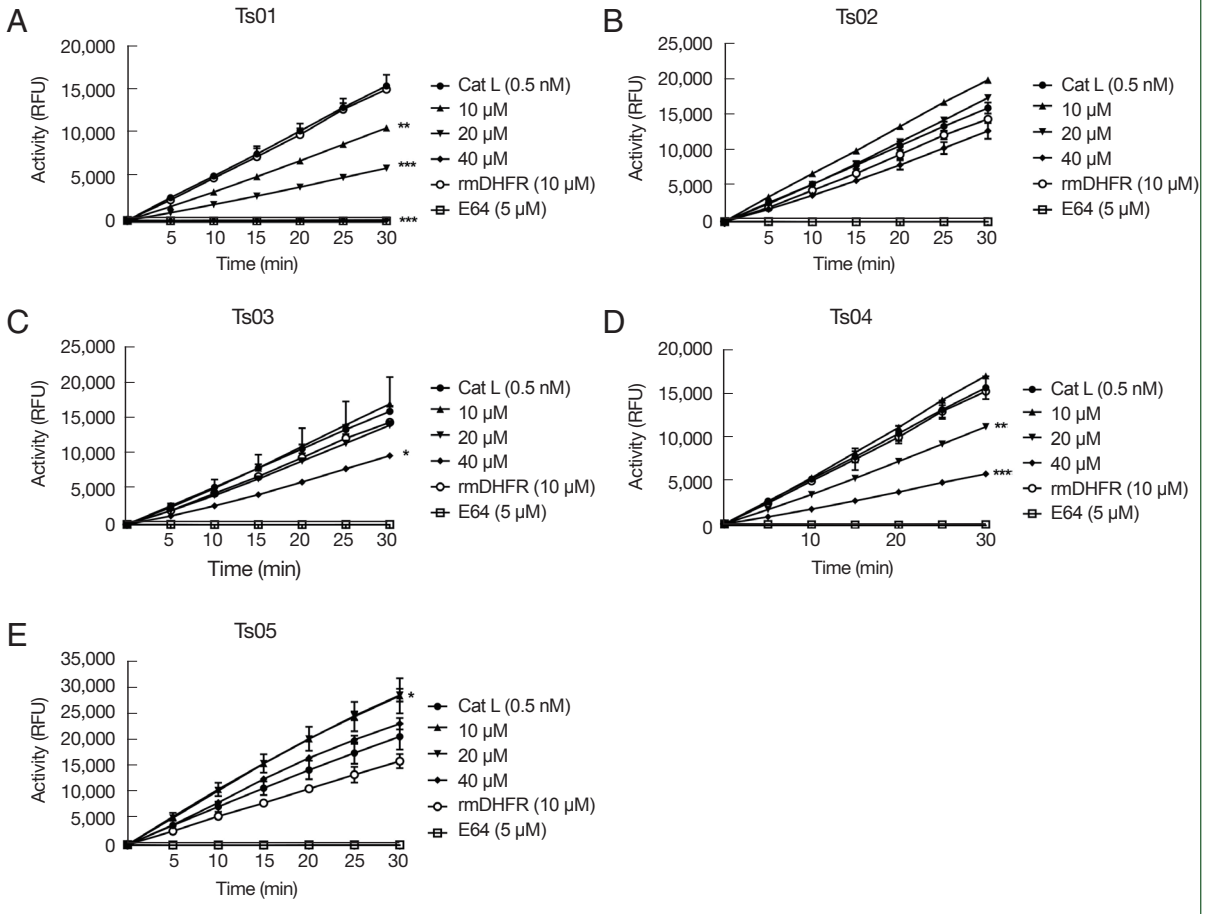

Fig. 2. Cathepsin L (CatL) Inhibitory activity of rTsCstN: (A) Ts01, (B) Ts02, (C) Ts03, (D) Ts04, and (E) Ts05 at 10, 20, and 40 μM. rmDHFR served as an irrelevant control while E64 functioned as a positive inhibitor control. Results were expressed as the mean ± SD. Experiments were performed in duplicate with two independent experiments. One-way ANOVA test was used for analysis: *$P < 0.05$, **$P < 0.001$, and ***$P < 0.0001$ indicate a significant difference between CatL activity only and CatL activity inhibited by rTsCstN, rTs02, rTs03, rTs04, and rTs05.

the protein concentrations were measured. Western blot analysis of rTs01, rTs02, rTs03, and rTs04 with mouse anti-His Tag antibodies confirmed that the antibody reacted with all recombinant proteins (Fig. 1B). rmDHFR was used as an irrelevant control protein and was expressed in a soluble form at a molecular weights of approximately 24 kDa and subsequently purified under native conditions (Supplementary Fig. S8).

Evaluation of the protease inhibitory activity of TsCstN and truncates

To evaluate the inhibitory properties of TsCstN truncates against CatL, different concentrations of rTs01 and truncates (10, 20, and 40 μM) were incubated with 0.5 nM CatL before

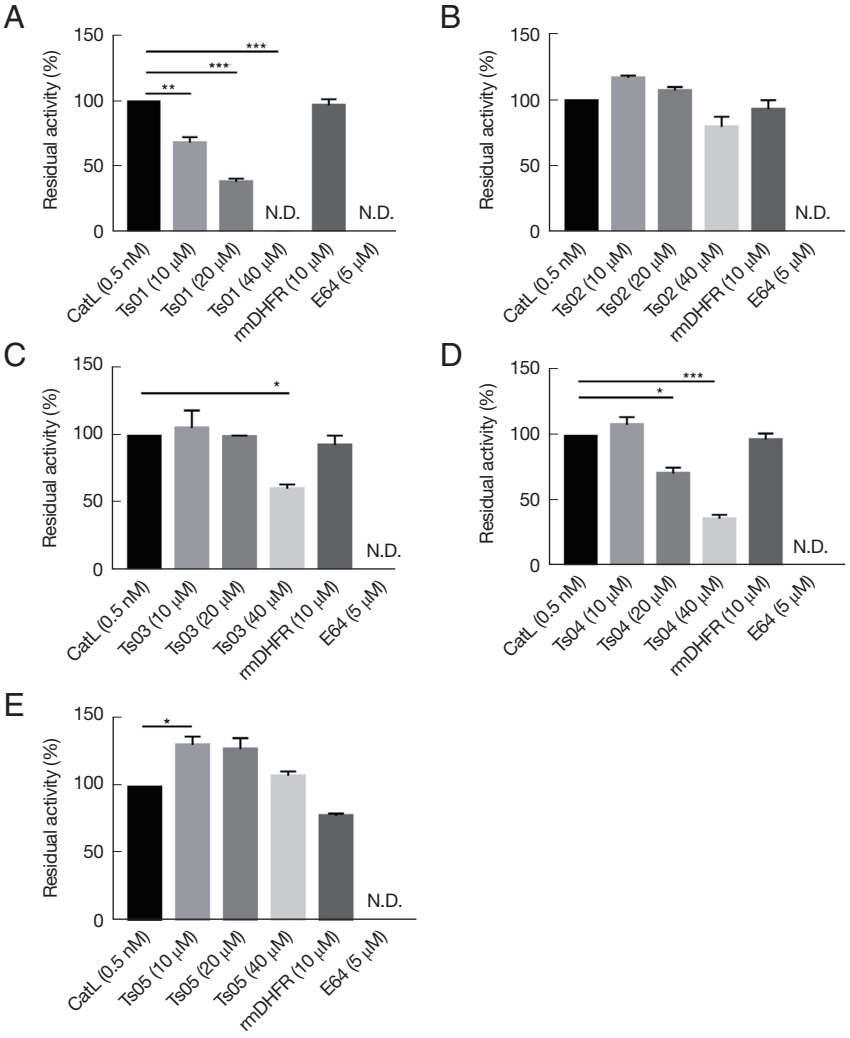

Fig. 3. Residual activity (%) of cathepsin L (CatL) when inhibited by TsCstN (Ts01) or truncates (Ts02, Ts03, Ts04, and Ts05) in various concentrations at 10, 20, and 40 μM. rmDHFR served as an irrelevant control, whereas E64 functioned as a positive inhibitor control. Results are expressed as the mean ± SD. One-way ANOVA test was used for analysis: *$P < 0.05$, **$P < 0.001$, and ***$P < 0.0001$ indicate a significant difference between CatL activity only and CatL activity inhibited by rTsCstN, rTs02, rTs03, rTs04, and rTs05. N.D., not detectable.

reacting with a CatL-specific fluorogenic substrate. rTs01 significantly reduced CatL activity at concentrations of 10 μM ($P < 0.001$), 20 μM ($P < 0.0001$), and 40 μM ($P < 0.0001$), with complete inhibition observed at 40 μM compared with CatL alone (Fig. 2A). Similarly, rTs04 significantly reduced CatL activity at 20 μM ($P < 0.001$) and 40 μM ($P < 0.0001$) compared with CatL alone (Fig. 2D), whereas rTs03 only inhibited CatL activity at 40 μM ($P < 0.05$) compared with CatL alone (Fig. 2C). CatL activity was not inhibited by rTs02 (Fig. 2B). Surprisingly, rTs05 did not inhibit protease activity of CatL but did enhance protease activity (Fig. 2E). An irrelevant control protein (rmDHFR) did not suppress CatL activity while E64, a broad-spectrum cysteine protease inhibitor, completely inhibited CatL activity.

The percentage residual activity of CatL after treatment rTs01 and its truncates (at 10, 20, and 40 μM) at the endpoint is summarized in Fig. 3. These results align with the inhibitory effects observed in Fig. 2, where rTs01 and certain truncates demonstrated significant reductions in CatL activity.

Discussion

Cystatins derived from parasitic helminths have been characterized as immunomodulatory molecules that affect immune responses such as the suppression of T-cell proliferation, inhibition of antigen processing and presentation in antigen-presenting cells, and regulation of pattern recognition receptors, and particularly the TLR-4 receptor on macrophages and dendritic cells, which release suppressive cytokines enabling long-term survival in their host [21]. Our earlier research showed that TsCstN derived from *T. spiralis* reduced CatL-associated inflammation [6]. In this study, we focused on characterizing protease inhibitory regions of TsCstN to modify them into a small peptide inhibitor against excessive inflammation.

The sequence analysis of TsCstN identified a signal peptide (1–20 aa), an N-glycosylation region, and two potential disulfide bonds, which are indicated as type 2 cystatins. In our previous study, TsCstN was initially identified as a hypothetical protein, necessitating structural annotation to elucidate its potential function. The simulated 3D structure had the closest homology with human cystatin E/M (HsCstM). However, multiple alignment comparisons of the protein sequence of TsCstN with those of HsCstM and a few selected cystatin homologs through show that TsCstN lacks the conserved regions typical of type 2 cystatins. These include the absence of the glycine residue at the N-terminus, the QXVXG motif at L1, and the PW motif at L2 [6]. In the present study, additional helminth cystatin homologs were incorporated to ascertain the presence of cystatin-conserved regions. Despite these efforts, the results remained consistent with our earlier findings, and crystallization may be necessary to determine the accurate 3D structure of TsCstN. This would provide insights into the biological function of the protein and its significance in *T. spiralis*, including its role in host-parasite interactions. In previous studies, hypothetical proteins have been crystallized from parasites such as *Giardia lamblia*, *Leishmania major*, and *Schistosoma mansoni* [22-24]. This approach is crucial for gaining deeper insights into their biology and for the development of effective drugs and vaccines.

To clarify the inhibitory reactive region of TsCstN, the protein truncated sites were de-

signed based on the 2D structure of TsCstN as predicted by five different programs, including Jpred4, PHD, Phyre2, PSIPRED, and PredictProtein. Impressively, all programs predicted a similar 2D structure of TsCstN, which comprised five reactive regions: α1, L1 (β1 and β2), irregular coil, L2 (β3 and β4), and α2. Four distinct truncates, Ts02–Ts05, were generated by deleting the protein regions in TsCstN. The 2D structures of all truncates, predicted by five different programs, showed consistent outcomes, indicating that the different deletions did not affect 2D structure conformation alterations.

I-TASSER and AlphaFold2 were used to simulate the 3D structures of Ts01 and truncates. AlphaFold2 could accurately predict the full-length and truncated TsCstN protein structures with a high degree of structural similarity to the hypothesized protein. Certain truncated structures, particularly Ts04, exhibited good quality with a high percentage of favored regions. I-TASSER is a widely used technique that creates precise 3D models of protein structures by combining several approaches, such as threading, ab initio modeling, and structural refining [17]. AlphaFold2 predicts the 3D structure of proteins by using a deep neural network architecture and advanced machine learning techniques that utilize amino acid sequences [18]. Consequently AlphaFold2 can predict 3D structures with high accuracy and prediction compared with I-TASSER. AlphaFold, an artificial intelligence-based technique for predicting protein structures, has demonstrated remarkable accuracy in model prediction and can consistently predict protein structures with atomic accuracy, even when no comparable structure is available [18]. Recently, AlphaFold has been extensively used in several structural biology fields, such as drug discovery, protein design, and protein function prediction [25-27].

The inhibitory properties of the recombinant TsCstN and truncates were confirmed in reactions against cathepsin L. Cysteine proteases, including CatB, CatL, and CatS, are involved in antigen presentation processing. CatB processes the antigen, and CatL and CatS degrade the major histocompatibility complex (MHC) invariant chain (Ii) in antigen processing cells, activating MHC-II formation [28,29]. Several studies have shown that parasite cystatins can inhibit cathepsin during antigen presentation [21]. In our previous study, rTsCstN specifically inhibited the activity of CatL but not that of CatB or CatS [6]. Therefore, this study focused on determining the inhibitory function of truncates against CatL. The 3D structure of TsCstN and other cystatins contains the NH2-terminal α1, L1, and L2, structures which are necessary to interact with the active site of cysteine proteases for inhibition [6,30]. rTs01 and rTs04 effectively inhibited CatL compared with other truncates. Thus, α1 and L1 are essential in inhibiting CatL activity, although rTs04 was less effective in reducing CatL activity than TsCstN-WT at the same dose. In human cystatin A, the NH2-terminal α1 region is involved in cysteine protease activity, and different NH2-terminal truncations cause the loss of inhibitory activity [31]. An important role of the α1 region of TsCstN in CatL inhibition was suggested by Ts02, which lacks α1 but contains L1 and L2. This truncate lost the ability to inhibit CatL. However, the α1 alone in Ts05 failed to inhibit CatL. Interestingly, rTs05 unexpectedly increased CatL activity instead of inhibiting it. However, the cause of this increase remains unknown and has not been addressed elsewhere and will therefore be explored in future study. We conclude that the α1 and L1 regions (Fig. 4), comprising 52 amino acid residues (DLSELDEAKNYIYQSDLQTGRGNFRKVLKVRNVDTSDGLSLTIDALPTTCPV) (Supplementary Table S3) are sufficient to

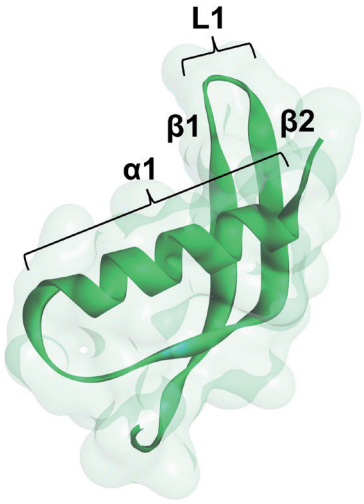

Fig. 4. Essential cathepsin L (CatL) inhibitory domain of Ts04, consisting of α1, β1, and β2. The interaction between β1 and β2 forms the loop 1 (L1) structure. Both α1 and L1 are necessary for inhibiting CatL activity.

inhibit CatL, whereas L2 helps increase CatL inhibitory activity.

Importantly, while TsCstN and its truncates exhibit CatL inhibitory activity, high concentrations of these recombinant proteins are required for effective inhibition. This could be due to several factors, including TsCstN not being a specific inhibitor for CatL or potential conformational inaccuracies when expressed in a prokaryotic system. Therefore, native TsCstN must be purified to test its protease inhibitory activity. Additionally, optimizing the expression system and conducting crystallization studies should be prioritized in future work.

In conclusion, CatL activity was efficiently inhibited by TsCstN truncates, particularly Ts04, which contains α1 and L1 regions. Further research should be conducted regarding the ex vivo reduction of macrophage inflammation and in vivo anti-inflammation in animal models. Furthermore, protein dynamics simulations, inhibitor types, enzyme kinetics, and molecular docking techniques can be used to deepen our understanding of the function of these proteins.

Acknowledgements

This research project was supported by Mahidol University under the New Discovery and Frontier Research grant for Fiscal Year 2021 through Poom Adisakwattana (NDFR 03/2564).

References

1. Murrell KD, Pozio E. Worldwide occurrence and impact of human trichinellosis, 1986-2009. *Emerg Infect Dis* 2011;17(12):2194-2202. https://doi.org/10.3201/eid1712.110896
2. Despommier DD. How does *Trichinella spiralis* make itself at home? *Parasitol Today* 1998;14(8):318-323. https://doi.org/10.1016/s0169-4758(98)01287-3
3. Fabre MV, Beiting DP, Bliss SK, Appleton JA. Immunity to *Trichinella spiralis* muscle infection. *Vet Parasitol* 2009;159(3-4):245-248. https://doi.org/10.1016/j.vetpar.2008.10.051
4. Alvarez JI. Inhibition of toll like receptor immune responses by microbial pathogens. *Front Biosci* 2005;10:582-587. https://doi.org/10.2741/1554
5. Aranzamendi C, Fransen F, Langelaar M, Franssen F, van der Ley P, et al. *Trichinella spiralis*-secreted products modulate DC functionality and expand regulatory T cells *in vitro*. *Parasite Immunol* 2012;34(4):210-223. https://doi.org/10.1111/j.1365-3024.2012.01353.x
6. Kobpornchai P, Flynn RJ, Reamtong O, Rittisoonthorn N, Kosoltanapiwat N, et al. A novel cystatin derived from *Trichinella spiralis* suppresses macrophage-mediated inflammatory responses. *PLoS Negl Trop Dis* 2020;14(4):e0008192. https://doi.org/10.1371/journal.pntd.0008192
7. Kobpornchai P, Tiffney EA, Adisakwattana P, Flynn RJ. *Trichinella spiralis* cystatin, TsCstN, modulates STAT4/IL-12 to specifically suppress IFN-γ production. *Cell Immunol* 2021;362:104303. https://doi.org/10.1016/j.cellimm.2021.104303
8. Khatri V, Chauhan N, Prasanna Kumar SB, Kalyanasundaram R. Peptide fragments of cystatin protein from filarial parasite has potent anti-inflammatory effect on DSS-induced colitis in mouse. *J Immunol* 2020;204(Suppl):237.28. https://doi.org/10.4049/jimmunol.204.Supp.237.28
9. Pikuła M, Smużyńska M, Krzystyniak A, Zieliński M, Langa P, et al. Cystatin C peptidomimetic derivative with antimicrobial properties as a potential compound against wound infections. *Bioorg Med Chem* 2017;25(4):1431-1439. https://doi.org/10.1016/j.bmc.2017.01.004
10. Hill M, Cox J. Cystatin C peptide effects on B16F10 melanoma cells. *CellBio* 2021;10:1-9. https://doi.org/10.4236/cellbio.2021.101001
11. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. Clustal W and Clustal X version 2.0. *Bioinformatics* 2007;23(21):2947-2948. https://doi.org/10.1093/bioinformatics/btm404
12. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics* 2000;16(4):404-405. https://doi.org/10.1093/bioinformatics/16.4.404
13. Drozdetskiy A, Cole C, Procter J, Barton GJ. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* 2015;43(W1):W389-394. doi: 10.1093/nar/gkv332
14. Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol* 1993;232(2):584-599. https://doi.org/10.1006/jmbi.1993.1413
15. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 2015;10(6):845-858. https://doi.org/10.1038/nprot.2015.053
16. Hall TA. Bioedit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/ NT. *Nucleic Acid Symposium Series* 1999;41:95-98.
17. Yang J, Yan R, Roy A, Xu D, Poisson J, et al. The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 2015;12(1):7-8. https://doi.org/10.1038/nmeth.3213
18. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;596(7873):583-589. https://doi.org/10.1038/s41586-021-03819-2
19. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 1993;26(2):283-291. https://doi.org/10.1107/S0021889892009944
20. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 2018;46(W1):296-303. https://doi.org/10.1093/nar/gky427
21. Khatri V, Chauhan N, Kalyanasundaram R. Parasite cystatin: immunomodulatory molecule with therapeutic activity against immune mediated disorders. *Pathogens* 2020;9(6):431. https://doi.org/10.3390/pathogens9060431
22. Beard DK, Bristol S, Cosby K, Davis A, Manning C, et al. Crystal structure of a hypothetical protein from *Giardia lamblia*. Corrigendum. *Acta Crystallogr F Struct Biol Commun* 2022;78(3):143. https://doi.10.1107/S2053230X22001704
23. Arakaki T, Le Trong I, Phizicky E, Quartley E, DeTitta G, et al. Structure of Lmaj006129AAA, a hypothetical protein from *Leishmania major*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2006;62(3):175-179. https://doi.10.1107/S1744309106005902
24. Masamba P, Weber BW, Sewell BT, Kappo AP. Crystallization and preliminary structural determination of the universal stress G4LZI3 protein from *Schistosoma mansoni*. *Inform Med Unlocked* 2022;32:101057. https://doi.org/10.1016/j.imu.2022.101057
25. Yang Z, Zeng X, Zhao Y, Chen R. AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduct Target Ther* 2023;8(1):115. https://doi.org/10.1038/s41392-023-01381-z
26. Ros-Lucas A, Martinez-Peinado N, Bastida J, Gascón J, Alonso-Padilla J. The use of AlphaFold for *in silico* exploration of drug targets in the parasite *Trypanosoma cruzi*. *Front Cell Infect Microbiol* 2022;12:944748. https://doi.org/10.3389/fcimb.2022.944748

27. Pak MA, Markhieva KA, Novikova MS, Petrov DS, Vorobyev IS, et al. Using AlphaFold to predict the impact of single mutations on protein stability and function. *PLoS One* 2023;18(3): e0282689. https://doi.org/10.1371/journal.pone.0282689
28. Honey K, Rudensky AY. Lysosomal cysteine proteases regulate antigen presentation. *Nat Rev Immunol* 2003;3(6):472-482. https://doi.org/10.1038/nri1110
29. Zhang T, Maekawa Y, Hanba J, Dainichi T, Nashed BF, et al. Lysosomal cathepsin B plays an important role in antigen processing, while cathepsin D is involved in degradation of the invariant chain inovalbumin-immunized mice. *Immunology* 2000;100(1):13-20. https://doi.org/10.1046/j.1365-2567.2000.00000.x
30. Salát J, Paesen GC, Rezácová P, Kotsyfakis M, Kovárová Z, et al. Crystal structure and functional characterization of an immunomodulatory salivary cystatin from the soft tick Ornithodoros moubata. *Biochem J* 2010;429(1):103-112. https://doi.org/10.1042/BJ20100280
31. Shibuya K, Kaji H, Itoh T, Ohyama Y, Tsujikami A, et al. Human cystatin A is inactivated by engineered truncation. The NH2-terminal region of the cysteine proteinase inhibitor is essential for expression of its inhibitory activity. *Biochemistry* 1995;34(38):12185-12192. https://doi.org/10.1021/bi00038a012