Is the absence of alpha-helix 2 in the appendant structure region the major contributor to structural instability of human cystatin C?

Xuejie Zhou1*, Xian Lu1*, Shuzhen Qin1*, Linan Xu2, Xiaoying Chong1, Junqing Liu1, Pingyu Yan1, Rui Sun1, Ian P. Hurley3, Gary W Jones3, Wang Qiuyu1* and Jianwei He1*

1School of Life Science, Liaoning University, Shenyang 110036, China;
2Department of Biology, Maynooth University, Maynooth, Co. Kildare, Ireland;
3Centre for Biomedical Science Research, School of Clinical and Applied Sciences, Faculty of Health and Social Sciences, Leeds Beckett University, City Campus, Leeds LS1 3HE, United Kingdom

*These authors contributed equally to this work.

*Corresponding author. Email: jwhe@lnu.edu.cn; qiuyuwang@lnu.edu.cn

Acknowledgements

This work was supported by grants from National Natural Science Foundation of China (No. 31670103) and was partially sponsored by Key Laboratory Projects of Liaoning Province (LZ2015044).
Key words: Cystatin,Appending Sequence, α-helix 2, Molecular dynamics, Stability

List of Abbreviations:

- AS: appending structure
- cC: chicken cystatin
- HCC: human cystatin C
- MD: molecular dynamics
- PVDF: polyvinylidene difluoride
- RMSD: root-mean-square deviation
- RMSF: root-mean-square fluctuation
- WT: wild type

Introduction
Human cystatin C (HCC) is a 120-amino acid protein that belongs to the cystatin super-family. It is expressed in all nucleated human cells and is found in almost all tissues and body fluids. HCC plays a physiological role by acting as a high-affinity inhibitor of cathepsins B, H, K, L and S. It also acts as a potent cysteine protease inhibitor, exhibiting pleiotropic roles in human vascular pathophysiology and serving as a biomarker for kidney disease. HCC is present in high concentrations in cerebrospinal fluid. Abnormal changes in the HCC expression in the brain can lead to various neurological disorders and neurodegenerative diseases, such as human cystatin C amyloid angiopathy and recurrent hemorrhagic stroke, respectively. However, wild-type (WT) HCC can also form part of the amyloid deposits in the brain arteries of elderly people with amyloid angiopathy.

Although the multiple roles of HCC have been broadly investigated, the most well-characterized cysteine protease inhibitor in cystatin type 2 super family is chicken cystatin (cC) because of its thermophilic and pH stabilities. Chicken cystatin is a secreted protein of low molecular weight (~14kDa). The protein was originally isolated from hen egg-white, and subsequently found in various chicken tissues. The secondary structure of cC consists of a five-stranded antiparallel β-sheet and a large central α-helix. The connectivity within the β-strands is: (N)-β1-(α)-β2-Loop1-β3-(AS)-β4-Loop2-β5-(C), in which a region consisting of 20 amino acid residues, called the appending structure (AS) exists between the β3-strands and β4-strands (Fig. 1) (He et al., 2013).

NMR studies have confirmed that the 3-D structure of the cC monomer is similar to that described for the monomeric part of HCC (Ekiel et al., 1997), with the structural similarity being 62.5%. In addition, the two proteins also share similar physicochemical properties (Yu et al., 2010). Nevertheless, the heat stability of cC is much higher than that of HCC (Zerovnik, Cimerman, Kos, Turk, & Lohner, 1997). The main difference between the two structures is the presence/absence of α-helix2 in AS (Fig. 1A). In our previous work, we have demonstrated that AS plays an important role in amyloidogenic cystatin dimerization between two cystatin monomers (Yu et al., 2012). Alignment of HCC and cC showed that E82 of cC corresponds to P84 of HCC (Fig. 1B). HCC does
not contain \( \alpha \)-helix2 in the AS region, but P84 is located in the center of the AS region. Proline is known to act as a structural disruptor if it is present in the middle of a regular secondary structure element such as \( \alpha \)-helices and \( \beta \)-sheets (MacArthur & Thornton, 1991). We speculated that the P84 may disrupt the \( \alpha \)-helix2 in HCC so that the HCC monomer is less stable than that of cC, and this could eventually lead to the dimerization and further fibrillization of HCC. To clearly demonstrate this hypothesis, we constructed an AS truncated mutant (\( \Delta W \)) in which residues 77-85 of \( \alpha \)-helix2 were deleted. Moreover, the disulfide bond formed between Cys71-Cys81 might be a key factor in initiating the unfolding of cC (Cys73-Cys83 in HCC) (Kolodziejczyk et al., 2010). To minimize the effect induced by the loss of the Cys71-Cys81 disulfide bond and to investigate the importance of the proline residue in HCC, a E82P mutant of cC was employed in the study.

Although cC has been used as an alternative \textit{in vitro} model to study the amyloidogenesis of HCC, it is still not known whether the large difference in thermodynamic stability between HCC and cC is actually caused by the only structural difference between the two proteins (Ekiel et al., 1997). In this study, we evaluated the effect of \( \alpha \)-helix2 on the structural stability of cC by comparing the intracellular and extracellular (secreted) levels of WT cC, E82P, \( \Delta W \) and a classic amyloid mutant (I66Q) in \textit{Pichia pastoris} X-33 (\textit{P. pastoris} X-33). The results suggested that the absence of \( \alpha \)-helix 2 in the AS region could induce structural instability in cC, but further investigation is needed to determine what kind of structural change can trigger the process of dimerization and fibrillization. Our experimental and \textit{in silico} studies did determine the influence displayed by the structural difference between HCC and cC, and this could shed more light on the mechanism of cystatin-associated amyloidosis.
Figure 1. Comparison of the structures of human and chicken cystatins. (A) The three-dimensional structures of the human cystatin C (HCC) and chicken cystatin C (cC) (Bode et al., 1988; Kolodziejczyk et al., 2010). (B) Comparison of the amino acid sequences that span β3 and β4 of HCC and cC.

Materials and methods

Molecular Dynamic Simulations

Molecular dynamics (MD) simulations were carried out using the GROMACS 4.5.0 software package. The structure of cC (PDB code: 1CEW) was used as the wild type (WT), whereas I66Q, ΔW and E82P models were constructed from WT using Swiss-PDB Viewer (Guex & Peitsch, 1997). The GROMOS96 43al force field (Pronk et al., 2013) was applied in all simulations. The models were separately immersed in the cubic boxes with a distance between protein and box edges of at least 10Å. The linear constraint solver (LINCS) method was used to constrain bond lengths, allowing an integration step of 2 fs. Electrostatic interactions were calculated using the particle mesh Ewald algorithm with a distance cut off of 9 Å. Lennard-Jones potential was used for describing the short-range attractive and repulsive dispersion interactions with a 10-Å cut off. The temperature and pressure were coupled using V-rescale and Parrinello-Rahman algorithm respectively (Yu et al., 2010). The system was neutralized by adding
the counter ions Cl\(^{-}\), and then energy-minimized using the Steepest Descent algorithm to remove the inappropriate contacts. After equilibration, MD simulations were carried out over a 20-ns period at 330 K, pH2.0 and 1 bar pressure. The coordinate trajectories were saved every 1 ps for subsequent data analysis.

**Strains, Plasmids and culture conditions**

*P. pastoris* X-33 and Plasmids pPICZ\(\alpha\)A were purchased from Invitrogen. Yeast cells were cultured at 30\(^{\circ}\)C in Yeast Peptone Dextrose medium (YPD) (1% yeast extract, 2% peptone and 2% glucose) and recombinant proteins were induced in Yeast Extract Peptone Medium (YPM) (1% yeast extract, 2% peptone, 0.5% Methanol (v/v)). Recombinant DNA manipulations were carried out in *E. coli* strain TOP10 (Invitrogen), which was cultured at 37\(^{\circ}\)C in LB medium. pPICZ\(\alpha\)A-WT and pPICZ\(\alpha\)A-I66Q were constructed, and *E. coli* and *P. pastoris* strain were transformed with these plasmids as described in our previous work (He et al., 2005).

**Construction of expression plasmids**

The truncated mutant (\(\Delta W\)) which carried a deletion at amino acid residues 77-85 was amplified from the DNA of WT eC using overlap extension PCR. Two pairs of primers were used: the first pair contained the forward primer 5’-GGGCTCGAGAAAAAGAGCGAGGA-3’ (\(XhoI\) site underlined), and the reverse primer, 5’-ACCTGATGACTTGGGGCAAGTTGT-3’, whereas the second pair contained the forward primer 5’-GCACAACTTGCCCCAAAGTCATCAGGTGAGCCAGAGATGGCTAAGTATACC-3’ (Overlapping section underlined) 5 and the reverse primer 5’-GGGCTCTAGATTACTGGCACTTGCT-3’ (\(XbaI\) site underlined) (Hussain & Chong, 2016). Similarly, the E82P substitution was introduced using the forward primer 5’-CAGAGCTGCCCATTCACGAT-3’ and the reverse primer 5’-GTGGAATGGGCAGCTCCAGGAG-3’ (mutation site underlined). The PCR product was digested with *Xho* I and *Xba* I and ligated into *Xho* I-*Xba* I digested pPICZ\(\alpha\)A vector, generating pPICZ\(\alpha\)A/\(\Delta W\) vector or pPICZ\(\alpha\)A/E82P vectors.
**Transformation into P. pastoris and small-scale expression of cC mutants**

*Pichia pastoris* X33 was transformed with *Sac* I-linearized pPICZαA-ΔW/ pPICZαA-E82P by electroporation and positive transformants were selected on YPDS plates (1% yeast extract, 2% peptone, 2% glucose, 1M sorbitol and 2% agar) containing 100 µg/mL zeocin (Invitrogen). To induce the expression of the recombinant protein, *P. pastoris* cells were first grown in 25 ml YPD medium (1% yeast extract, 2% peptone and 2% glucose) at 30°C for about 24 h with shaking (200 rpm) until the OD600 was about 5.0. The cells were harvested and resuspended in 50 ml YPM medium (1% yeast extract, 2% peptone) and incubated for 72 h. The culture was supplemented with 0.5% (v/v) methanol at every 24 h.

**Preparation of extracellular and intracellular protein sample**

For extracellular protein sample preparation, yeast cultures were centrifuged at 4°C 1,900× g for 10 min to pellet the cells, and the supernatant was further centrifuged at 4°C 4,866 × g for 30 min to remove the insoluble protein aggregate. The supernatant was used for extracellular protein analysis.

For intracellular protein sample preparation, yeast cultures were centrifuged at 4°C 1,900 × g for 10 min. The cell pellets were resuspended in lysis buffer (10 mM Tris-HCl pH 7.4, 0.1 M NaCl, 1 mM EDTA, 10% SDS, 4% glycerin (v/v), 1% Triton X-100, 5mM DTT and 1mM PMSF), and then mixed with Zirconium beads and shaken violently to break up the cells for five times every 30s. The cell extracts were centrifuged at 4°C 12,848 × g for 10 min and the supernatants, which contained intracellular proteins, were retained. The protein concentration in each supernatant was determined by BCA protein assay kit (Beyotime, China) and normalized before the analysis. All of the protein samples were treated with protein loading buffer at 100°C for 5 min and stored at -80°C in preparation for SDS polyacrylamide gel electrophoresis (PAGE) and Western blotting analysis.

**SDS-PAGE and Western blotting**
Protein samples were analyzed by 15% SDS-PAGE under denaturing conditions. Proteins in the gels were visualized after staining with Coomassie Brilliant Blue. After electrophoresis, the protein bands were transferred to a polyvinylidene difluoride (PVDF) membrane for 1 h 20 min at 100 mA in blotting buffer (20 mM glycine, 25 mM Tris, 20% methanol). The membrane was blocked by incubation in Tris-buffered saline (TBS) containing 5% BSA and 0.05% Tween 20 at room temperature for 3 h on an orbital shaker. This was followed by incubation with rabbit anti-chicken cystatin antiserum (2 × 10⁻³ in 0.05% Tween 20) in Tween-TBS at room temperature for 1 h. After that, the membrane was extensively washed and then incubated with anti–rabbit peroxidase conjugate (1 × 10⁻⁴ in Tween-TBS) at room temperature for 1 h and 20 min, followed by detection performed with an ECL-Plus kit (Beyotime, China).

Results and Discussions

The α-helix 2 contributes to the stabilization of chicken Cystatin C

Our previous report has suggested that AS might trigger the unfolding of cC and accelerates the formation of cystatin aggregation or amyloid fibril. Comparison of the crystal structure of cC with that of HCC showed the main difference between HCC and cC is the presence/absence of α-helix2 in AS (Fig. 1A). To investigate whether α-helix 2 would contribute to the difference in amyloid fibril formation between HCC and cC, we constructed a mutant form of cC (ΔW) lacking the α-helix 2 and performed molecular dynamic simulation to investigate the potential influence of the α-helix 2. Considering that α-helix 2 contributes to the formation of the C71-C81 disulfide bond and E82 of cC is replaced by a rigid and structural disruptor, P84, in HCC, we replaced E82 of cC with a proline to disrupt α-helix 2. This change would have minimal effect compared with the direct breakage of the C71-C81 disulfide bond.

To predict the stability of the ΔW and E82P mutants, the root-mean-square deviations (RMSD) of the backbone α carbon atoms of cC were calculated for all simulations. It has been demonstrated that in MD simulations, high temperature and low pH value can accelerate protein unfolding without changing the pathway of the
unfolding process (Day, Bennion, Ham, & Daggett, 2002). Therefore, we performed the simulations under the same extreme conditions (330 K (57 °C) and pH 2) as those used in our previous experiments (He et al., 2005; Yu et al., 2012; Yu et al., 2010), which allow the system to reach the equilibrated regions in minimum simulation time and with minimum computational expense. The RMSD value reached a relative equilibrium region in 10 ns in all simulations (Fig. 2). As previously reported, the classic amyloidogenic mutant I66Q displays an obvious RMSD fluctuation compared to WT (Fig. 2), and this is considered to be a characteristic that promotes amyloid fibrillization (Yu et al., 2012). Interestingly, the RMSD of I66Q and ΔW increased dramatically as soon as the simulation began. However, the RMSD of E82P gradually increased during the first 7 ns, suggesting that I66Q and ΔW would be more likely to become unstable compared with WT and E82P. The lower stability of ΔW might be caused by the deletion of the 9 residues, which resulted in the AS being too short to provide stability for the remaining structure. Interestingly, both ΔW and E82P demonstrated a similar RMSD profile to that observed for I66Q, implying that the disruption or conformational changes of α-helix 2 may promote amyloid fibrillization or simply accelerate the protein aggregation by destabilizing the secondary structure of the protein (Fig. 2). Indeed, the extracellular level of recombinant ΔW and E82P induced with 0.5% methanol for 3 days was reduced to various extents (Fig. 3A), especially the secretion of ΔW and E82P, and was consistent with the MD prediction (Fig. 2). The E82 to P mutation within the α-helix 2 could indeed disrupt the stability of cC to some extent, but with less sensitivity compared to the disruption observed for ΔW. A possible reason for this could be that in the case of E82P, the proline substitution may only promote the unfolding of α-helix 2 and weaken the C71-C81 disulfide bond, whereas in ΔW, both α-helix 2 and the C71-C81 disulfide bond were completely destroyed.
Figure 2. RMSDs of the Cα-positions during the 20 ns simulations. The RMSDs of the wild-type (WT) cC and the classic amyloidogenic mutant I66Q were obtained from our earlier study (Yu et al., 2012).

**Dimerization characterization of the recombinant proteins**

The secreted proteins were further confirmed by western blot analysis (Fig. 3C). We have previously reported that I66Q can easily form dimers in the culture medium during methanol induction. However, no dimers were found in the extracellular proteins of either the ΔW or E82P mutant, indicating that these two mutants are not prone to form dimers, but they would aggregate/unfold as a result of the destruction of the secondary structure. Intracellular expression of recombinant cC and its mutants was also analyzed by western blot. The level of monomeric WT cC was highest among the four recombinant cCs, but the major form of I66Q was a dimer, and fewer monomers were found inside the yeast cells than in the extracellular environment. The intracellular levels of monomeric ΔW and E82P were also lower than that of WT cC, but no dimer was detected, suggesting that ΔW and E82P may only promote amorphous aggregation, but not amyloid fibrillization as in the case of I66Q. Unfortunately, further amyloid fibrillization and aggregation assay could not be performed *in vitro* as no stable monomeric ΔW and E82P were obtained after purification. On balance, it is very likely
that the structural changes caused by ΔW and E82P simply induce protein unfolding, which were then subjected to degradation through the ER-associated degradation pathway, leading to lower protein production.

Figure 3. Analysis of recombinant cC expressed in *P. pastoris*. (A) SDS-PAGE of cC secreted into the culture supernatant. Western blot analysis of cC from the culture supernatant (B) and from the cell extract (C).

**Conformational changes caused by the disruption of the α-helix 2**

To further explore the detailed structural changes occurring in the ΔW and E82P mutants of cC, we calculated the average root mean-square fluctuations (RMSF) values. During the 20 ns of the MD simulation, major fluctuations were found clustering around α-helix 2 (AS region) for all three mutants (Fig. 4A). Fluctuation also occurred in α-helix 1 to some extent. The AS fluctuation displayed by ΔW was directly caused by the deletion of α-helix 2, which destroyed both Cys71-Cys81 and α-helix 2 (Fig. 4B). Notably, α-helix 2 in E82P was unfolded (Fig. 4B) as we predicted, supporting the hypothesis that P84 disrupts α-helix 2 in HCC and then decreases the stability of the HCC monomer, leading to dimerization of HCC and further fibrillization much more readily compared with monomeric cC. Helix 1 did not display obvious unfolding in ΔW
and the other mutants (Fig. 4B). Thus, the fluctuation around helix 1 might be caused by its own realignment. A previous report has suggested that the rotation of AS is related to cC fibrilization (Yu et al., 2012), and therefore, rotations caused by the absence of the Cys71-Cys81 disulfide bond and α-helix 2 in ΔW or E82P could further underline the significance of α-helix 2 in the conformational stability of cC.

Figure 4. (A) Averaged Cα RMSFs per residue in water during the convergence period of MD simulations. (B) Superposition of various states of WT (cyan) and I66Q (orange), ΔW (red) and E82P (blue) at the end of the 20 ns MD simulation respectively.

The disruption of α-helix 2 destabilizes the hydrophobic core

The interior hydrophobic core of cystatin provides the structural stability that favors the native state during the protein folding process (Szymanska et al., 2012). To further investigate whether there is a fluctuation in the hydrophobic core caused by mutations in α-helix 2 (ΔW and E82P), we analyzed the RMSDs of the hydrophobic core. I66Q could promote the expansion of the hydrophobic core (Fig. 5), consistent with previous report. In contrast, ΔW and E82P both exhibited obvious fluctuations
during the simulations, suggesting that the structural stability of cC was compromised to various degrees by the disruption of α-helix 2 (Fig. 5). Considering that α-helix 1 could contribute to formation of the hydrophobic core, the RMSF fluctuation of the α-helix 1 observed (Fig. 4A) might be caused by the extension of the hydrophobic core.

Figure 5. RMSDs of the hydrophobic-core backbone during the 20 ns simulations. The RMSDs of the wild-type (WT) cC and the classic amyloidogenic mutant I66Q were carried out in our earlier study (Yu et al., 2012).

To date, despite extensive investigation into the amyloidogenic properties of the two cystatins, little has been reported on whether the only structural difference between HCC and cC is the key factor causing the big difference in thermodynamic stability and dimerization difficulty between HCC and cC. Our results suggested that the absence of α-helix 2 in the AS region may be one of the contributing factors in the structural instability of human cystatin C. However, this change might not be the key factor that triggers the dimerization process, which would eventually lead to amyloid fibrillation. Other factors might also influence the dimerization and fibrillation of cystatins, such as mutations at “hot spot” V57 and L68 in the HCC sequence, hinge regions,
glycosylation, temperature and pH (He et al., 2006; Szymanska et al., 2012). Thus the detailed dimerization mechanism is still unclear and requires further investigation.

References

Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., . . . Turk, V. (1988). The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J*, 7(8), 2593-2599

Day, R., Bennion, B. J., Ham, S., & Daggett, V. (2002). Increasing temperature accelerates protein unfolding without changing the pathway of unfolding. *J Mol Biol*, 322(1), 189-203

Ekiel, I., Abrahamson, M., Fulton, D. B., Lindahl, P., Storer, A. C., Levadoux, W., . . . Gehring, K. (1997). NMR structural studies of human cystatin C dimers and monomers. *J Mol Biol*, 271(2), 266-277. doi: S0022-2836(97)91150-5 [pii]

10.1006/jmbi.1997.1150

Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, 18(15), 2714-2723. doi: 10.1002/elps.1150181505

He, J., Song, Y., Ueyama, N., Harada, A., Azakami, H., & Kato, A. (2005). Characterization of recombinant amyloidogenic chicken cystatin mutant I66Q expressed in yeast. *J Biochem*, 137(4), 477-485. doi: 10.1093/jb/mvi064

He, J., Song, Y., Ueyama, N., Saito, A., Azakami, H., & Kato, A. (2006). Prevention of amyloid fibril formation of amyloidogenic chicken cystatin by site-specific glycosylation in yeast. *Protein Sci*, 15(2), 213-222. doi: 15/2/213 [pii]

10.1100/ps.05175306

He, J., Xu, L., Zou, Z., Ueyama, N., Li, H., Kato, A., . . . Song, Y. (2013). Molecular dynamics simulation to investigate the impact of disulfide bond formation on conformational stability of chicken cystatin I66Q mutant. *J Biomol Struct Dyn*, 31(10), 1101-1110. doi: 10.1080/07391102.2012.721498

Hussain, H., & Chong, N. F. (2016). Combined Overlap Extension PCR Method for Improved Site Directed Mutagenesis. *Biomed Res Int*, 2016, 8041532. doi: 10.1155/2016/8041532

Kołodziejczyk, R., Michalska, K., Hernandez-Santoyo, A., Wahlbom, M., Grubb, A., & Jaskolski, M. (2010). Crystal structure of human cystatin C stabilized against amyloid formation. *FEBS J*, 277(7), 1726-1737. doi: 10.1111/j.1742-4658.2010.07596.x

MacArthur, M. W., & Thornton, J. M. (1991). Influence of proline residues on protein conformation. *J Mol Biol*, 218(2), 397-412

Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., . . . Lindahl, E. (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29(7), 845-854. doi: 10.1093/bioinformatics/btt055

Szymanska, A., Jankowska, E., Orlikowska, M., Behrendt, I., Czapiewska, P., & Rodziewicz-Motowidlo, S. (2012). Influence of point mutations on the stability, dimerization, and
oligomerization of human cystatin C and its L68Q variant. Front Mol Neurosci, 5, 82. doi: 10.3389/fnmol.2012.00082

Yu, Y., Liu, X., He, J., Zhang, M., Li, H., Wei, D., & Song, Y. (2012). Appendant structure plays an important role in amyloidogenic cystatin dimerization prior to domain swapping. J Biomol Struct Dyn, 30(1), 102-112. doi: 10.1080/07391102.2012.674282

Yu, Y., Wang, Y., He, J., Liu, Y., Li, H., Zhang, H., & Song, Y. (2010). Structural and dynamic properties of a new amyloidogenic chicken cystatin mutant I108T. J Biomol Struct Dyn, 27(5), 641-649. doi: c4299/Structural-and-Dynamic-Properties-of-a-New-Amyloidogenic-Chicken-Cystatin-Mutant-I108T-641-650-p17766.html [pii]

Zerovnik, E., Cimerman, N., Kos, J., Turk, V., & Lohner, K. (1997). Thermal denaturation of human cystatin C and two of its variants; comparison to chicken cystatin. Biol Chem, 378(10), 1199-1203