

Solution Structure of a Phytocystatin from *Ananas comosus* and Its Molecular Interaction with Papain

Deli Irene¹, Tse-Yu Chung², Bo-Jiun Chen¹, Ting-Hang Liu², Feng-Yin Li³, Jason T. C. Tzen², Cheng-I Wang⁴, Chia-Lin Chyan^{1*}

1 Department of Chemistry, National Dong Hwa University, Hualien, Taiwan, Republic of China, **2** Graduate Institute of Biotechnology, National Chung Hsing University, Taichung, Taiwan, Republic of China, **3** Department of Chemistry, National Chung Hsing University, Taichung, Taiwan, Republic of China, **4** Singapore Immunology Network, Biomedical Sciences Institutes, Agency for Science, Technology and Research (A*STAR), Singapore

Abstract

The structure of a recombinant pineapple cystatin (AcCYS) was determined by NMR with the RMSD of backbone and heavy atoms of twenty lowest energy structures of 0.56 and 1.11 Å, respectively. It reveals an unstructured N-terminal extension and a compact inhibitory domain comprising a four-stranded antiparallel β -sheet wrapped around a central α -helix. The three structural motifs (G⁴⁵, Q⁸⁹VXG, and W¹²⁰) putatively responsible for the interaction with papain-like proteases are located in one side of AcCYS. Significant chemical shift perturbations in two loop regions, residues 45 to 48 (GIYD) and residues 89 to 91 (QVV), of AcCYS strongly suggest their involvement in the binding to papain, consistent with studies on other members of the cystatin family. However, the highly conserved W120 appears not to be involved in the binding with papain as no chemical shift perturbation was observed. Chemical shift index analysis further indicates that the length of the α -helix is shortened upon association with papain. Collectively, our data suggest that AcCYS undergoes local secondary structural rearrangements when papain is brought into close contact. A molecular model of AcCYS/papain complex is proposed to illustrate the interaction between AcCYS and papain, indicating a complete blockade of the catalytic triad by AcCYS.

Citation: Irene D, Chung T-Y, Chen B-J, Liu T-H, Li F-Y, et al. (2012) Solution Structure of a Phytocystatin from *Ananas comosus* and Its Molecular Interaction with Papain. PLoS ONE 7(11): e47865. doi:10.1371/journal.pone.0047865

Editor: Annalisa Pastore, National Institute for Medical Research, Medical Research Council London, United Kingdom

Received: September 7, 2012; **Accepted:** September 24, 2012; **Published:** November 6, 2012

Copyright: © 2012 Irene et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study is supported by Taiwan National Science Council grant NSC-96-2113-M-259-005 to CLC. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: chyan@mail.ndhu.edu.tw

Introduction

Cystatins, the natural inhibitors of cysteine proteases are important regulatory proteins found in mammals, plants, and insects [1,2,3]. They are involved in various biological and pathological processes, such as tumor invasion, inflammation, antigen processing, dystrophy, and metastasis [4]. The cystatin superfamily in higher animal is generally classified into three subfamilies according to their size and the presence of internal disulfide bonds [5,6,7]. Members of family 1, termed stefins are small intracellular proteins of about 100 residues with no disulfide bond. Cystatins of family 2 are extracellular proteins of about 120 residues with at least two internal disulfide bonds. Members of family 3, termed kininogens, are relatively large blood-circulated glycoproteins composed of several family 2-like cystatin domains.

Previous studies have shown that the family 1 and 2 cystatin homologs share a common fold comprising an antiparallel β -sheet wrapped around a central α -helix [8,9,10,11,12]. The co-crystal structure of human stefin B/papain complex revealed the interaction of a wedge-shaped edge of the inhibitor with the enzyme active-site cleft [12]. Three structural elements essentially responsible for the specific binding of cystatins to the active-site cleft of papain-like cysteine proteases are a conserved Gly at the N-terminus, a QxVxG motif situated in the β -hairpin loop between the second and the third strand, and an aromatic residue on the β -

hairpin loop between the fourth and the fifth strand [13,14,15]. The parasite chagasins, cystatin-like proteins were newly found to have a similar recognition pattern with cysteine proteases [16,17,18,19,20]. Although chagasins strongly inhibit cysteine proteases, the three inhibitory loops of chagasins show low sequence homology to other cystatins.

Several plant cystatin genes were cloned and their deduced proteins were found homologous to animal cystatins in the past decade [21,22,23,24,25,26,27,28]. As a consequence of their sequence uniqueness, these plant cystatin-like proteins are classified as a new subfamily termed phytocystatins. Phytocystatins have been proposed to have several possible functions, including regulating the activity of endogenous cysteine proteases during different physiological processes such as seed maturation and germination, as well as responding to biotic and abiotic stresses [27,29,30,31]. Phytocystatins are also implicated in programmed cell death by modulating cysteine protease activities in the regulation of protein turnover [32,33]. They may also be involved in defense mechanisms to protect plants from the invasion of pathogens or the attack by pests [34]. Therefore, phytocystatins appear to function as inhibitors of both endogenous and exogenous cysteine proteases and may have potential applications in agriculture and medicine.

Sequence analysis suggests that phytocystatins also possess the three structural elements essential for the interaction with papain-

like proteases. In addition, phycocystatins contain a signature motif LARFAVXEHN in the α -helix forming sequence but lack disulfide bonds and potential glycosylation sites [6]. Although most of the phycocystatins are small proteins with molecular weight of 12–16 kDa, some phycocystatins possess an N-terminal signal sequence and/or a C-terminal extension that involves in the inhibition of a second family of cysteine proteases, legumain proteases [35]. Furthermore, several multicystatins that contain multiple copies of cystatin domains have also been identified [24,36,37,38].

A cDNA fragment encoding a phycocystatin of pineapple (*Ananas comosus* L.) stem was successfully cloned in our previous study [22]. To explore the potential applications of this pineapple cystatin (AcCYS), we aimed to unravel its inhibitory action at atomic level in this study. The recombinant AcCYS, consisting of 135 residues without any cysteine residue, was expressed in *Escherichia coli* and purified to apparent homogeneity. NMR techniques were employed to determine the structure of AcCYS and to characterize its interaction with papain. A docking model of AcCYS/papain was also proposed to illustrate the inhibitory action of AcCYS toward papain.

Materials and Methods

Expression and Purification of Full Length and Truncated AcCYS

The *AcCYS* gene encoding a cysteine protease inhibitor, cystatin was cloned from pineapple (*Ananas comosus*) stem [22]. The gene was then subcloned into a modified pET28a expression vector. A truncated form of *AcCYS* gene (*AcCYS-DL*) containing the coding gene of residue 28–135 was amplified by PCR, then subcloned into a modified pET28a expression vector. Both constructs were verified by DNA sequencing, and then transformed into *E. coli* BL21 (DE3) host for protein expression. After expression and cell disruption, the recombinant AcCYS (residue 1–135) and AcCYS_DL (residue 28–135) proteins were purified by anion exchange (DEAE, GE) followed by size exclusive (Superdex G75, GE) chromatography. ^{13}C , ^{15}N -labeled AcCYS and AcCYS_DL were overexpressed in a modified M9 medium containing 1 g/l $^{15}\text{NH}_4\text{Cl}$ and 2 g/l glucose ($^{13}\text{C}_6$ -glucose) as the sole ^{15}N and ^{13}C sources. The purified recombinant AcCYS and AcCYS_DL ran as a single band on SDS-PAGE. Protein concentrations of AcCYS and AcCYS_DL were determined by UV spectrometer using the extinction coefficient $\epsilon_{280\text{nm}}$, $18450 \text{ M}^{-1} \text{ cm}^{-1}$.

Preparation of NMR Samples and NMR Spectroscopy

The AcCYS and AcCYS_DL proteins were dissolved in a solution containing 20 mM KCl, 0.02% NaN_3 , and 10% (v/v) D_2O , and the pH was adjusted to 6.5. The final concentrations of the AcCYS were ranged from 0.7 to 1.1 mM for structure determination. AcCYS_DL/papain complex was prepared as follows: 1 mg ^{13}C , ^{15}N -labeled AcCYS_DL was dissolved into 1 ml solution containing 90% $\text{H}_2\text{O}/10\%$ D_2O , 20 mM KCl, 0.02% NaN_3 , at pH 6.5. The appropriate amount of unlabelled papain stock solution (0.5 mM) was added dropwise into the AcCYS_DL solution with gentle mixing to ensure AcCYS_DL/papain complex formation. The solution was then concentrated by centricon-30 ultrafiltration apparatus (Millipore Inc) to a final concentration of about 0.25 mM. The final NMR sample solutions were transferred to 5 mm Shigemi NMR tubes (Shigemi Co., Tokyo, Japan).

All NMR data were recorded at 303 K on Bruker AVANCE-600 spectrometer equipped with xyz-gradient TXI probe in

National Dong Hwa University. All the spectra were processed by TOPSPIN and analyzed by AURELIA on Linux workstations. Proton chemical shifts were referenced to DSS, and ^{13}C and ^{15}N chemical shifts were calibrated indirectly using the absolute frequency ratios.

Resonance Assignments and Structure Calculation of Free AcCYS

Resonance assignments of free AcCYS were obtained using a series of double- and triple-resonance NMR experiments and have been reported in our previous study [39]. To assign NOE-based distance restraints, 3D ^{13}C -separated and ^{15}N -separated NOESY-HSQC spectra were recorded on uniformly ^{13}C and ^{15}N -labeled samples respectively. The NOE crosspeaks were picked and quantified by peak-picking algorithm in Aurelia. The NOE peak intensities were converted into upper distance bounds using CALIB module in the torsional angle dynamics program CYANA [40]. The backbone dihedral angles (ϕ and ψ) restraints were derived by chemical shifts of $^1\text{H}^\alpha$, $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, ^{13}CO , and ^{15}N nuclei using the program TALOS [41]. Hydrogen bond restraints were assigned between slowly exchanging amide protons and their respective carbonyl acceptors deduced from the NOE data in combination with the secondary structure information predicted from CSI [42]. Initial structures were generated by the automated module, CANDID/NOEASSIGN in CYANA. These NOE assignments were carefully confirmed and erroneous ones corrected through examination of spectra. Additional NOE were then added manually before recalculation of structures by CYANA. The final 20 conformers with the lowest target function values were selected and further refined by the restrained simulated annealing and energy minimization algorithms in CNS 1.3 [43,44]. Graphical visualization and analyses of the structures were carried out with the programs MOLMOL [45] and PyMOL (DeLano Scientific). The geometric and stereochemical quality of the ensemble of structures was validated by PROCHECK-NMR [46].

Circular Dichroism (CD) Spectroscopy and Thermodynamic Stability Measurements

CD spectra were recorded on a Jasco J-715 circular dichroism spectrometer. CD spectra were collected using a cylindrical quartz cuvette with a 1 mm path-length. Measurements were made for the purified recombinant AcCYS and AcCYS-DL at a concentration of 10 μM . Each CD spectrum was averaged over 16 scans and corrected for the appropriate buffer baseline. All spectra are presented as the molar CD absorption coefficient ($\Delta\epsilon_M$). The contents of secondary structures were analyzed by the Dichroweb program [47,48].

The thermodynamic stability measurements for the unfolding of AcCYS and AcCYS_DL against denaturant GdnHCl was performed and calculated using a classical two-state model as shown in equation 1 [49].

$$\begin{aligned} \Delta G_D &= -RT \ln K = -RT \ln \left(\frac{f_U}{1-f_U} \right) \\ &= \Delta G_D^{H_2O} - m[\text{GdnHCl}] \end{aligned} \quad (1)$$

where ΔG_D and $\Delta G_D^{H_2O}$ represent the free energy of unfolding of proteins in the presence and absence of denaturant respectively; f_U is the fraction of unfolded state; m is a measure of the dependence of free energy on GdnHCl concentration. The fraction of the unfolded state (f_U) can be expressed as equation 2.

	10	20	30	40
AcCYS	MADEQPKKPVPA	AAEEAPAAEAE	AEAEAEGLHLEDD	QEPREHPIMGGI
Taro	-----	-----	-----	ALMGGI
Rice	-----	-----	-----	MSSDGGPVLGGV
Potato	-----	-----	-----	GI
CYS F	-----	-----	GPSPDTC	SQDLNSRVKPGFP
Chicken	-----	-----	-----	GAP
Stefin A	-----	-----	-----	MIPGGL
Stefin B	-----	-----	-----	MMSGAP
				**
	50	60	70	80
AcCYS	YDAPLN	NEN	GFDKEDL	ARFAVRE
Taro	VDVEG	-AQNS	AEVEELAR	FAVDEHNK
Rice	EPVG	--NEND	LHLVDLAR	FAVTEHNK
Potato	VNVP	--NPNN	TKFQELAR	FAIQDYNK
CYS F	KTIKT	---ND	PGVLQAARY	SVEKFNNCT
Chicken	VPVD	---ENDE	GLQRALQ	FAMAEYNRA
Stefin A	SEAKP	---ATPE	IQEIVDKV	KPQLEEKTN
Stefin B	SATQP	---ATA	ETQHIADQ	VRSQLEEKYN
				**
	100	110		
AcCYS	VSGMMHYLT	VEVND	-AGKKK	-----
Taro	VSGIMHHLT	VEVIE	-GGKKK	-----
Rice	VAGTLYYFT	TIEVKE	-GDAKK	-----
Potato	VAGIMYYIT	LAATD	DAGKKK	-----
CYS F	VKGLKYM	LEVEI	GRITTC	KKNQHLRLDD
Chicken	VSGIKYIL	QVEIG	RITTC	PKS-SGDLQ
Stefin A	VAGTNY	YIKV	RAGDNK	YMHLK
Stefin B	VAGTNY	FIKV	HVGDED	FVHLR

	120	130		
AcCYS	WEQVWMN	FRQLQ	EFTYLGDA	----
Taro	WVQAWLN	SKKLH	EFSPI	-----
Rice	WEKPWMD	FKELQ	EFKPVD	ASANA-
Potato	WVKEWED	FKKV	VEFKLV	-----
CYS F	WVVPWL	QHFEV	PVLRCH	HHHHH--
Chicken	YSIPWLN	QIKL	LESKCQ	-----
Stefin A	GQNE	DLVLT	GYQVD	KNKDEL
Stefin B	HENKPL	TL	SNYQT	NKAKHDEL
				*

Figure 1. Sequence alignment of AcCYS to the other cystatins. The sequence of AcCYS is compared with taro, rice, potato, human cystatin F, chicken egg white cystatin, human stefin A, and human stefin B. The characteristic motifs for putative interaction with target cysteine protease are marked by green asterisks. The unique motif, LARFAVxExN of phycocystatins is denoted by the red dots on top of the sequence. The α -helix and β -sheet regions are colored in red and blue, respectively.
doi:10.1371/journal.pone.0047865.g001

$$f_U = \frac{\exp\{-(\Delta G_D^{H_2O} - m[GdnHCl])/RT\}}{1 + \exp\{-(\Delta G_D^{H_2O} - m[GdnHCl])/RT\}} \quad (2)$$

Experimental data can be fitted according to this equation by the program Origin 6.0 (Microcal Software Inc.).

Inhibitory Activity of AcCYS and AcCYS-DL to Cysteine Proteases

The inhibition constants (K_i) for AcCYS and AcCYS_DL to a cysteine protease, papain (Sigma, Missouri, USA) were determined. A fluorogenic substrate, Z-Phe-Arg-7-amido-4-methylcoumarin hydrochloride, FRAMC (Calbiochem, Darmstadt, Germany) was used for the assay. The initial rates of cleavage of FRAMC by papain were obtained spectrofluorometrically. The inhibitory activity was assayed in a 100 mM sodium phosphate buffer, pH 6.3 with 10 mM EDTA, 400 mM NaCl, and 2 mM dithiothreitol. Substrate concentration was ranged from 0.3 to 5 μ M and substrate hydrolysis never exceeded 5%. The fluorescence was measured in an Amico-Bowman Series 2 spectrofluorometer with excitation and emission wavelengths at 346 and 450 nm, respectively. The kinetic data were processed by Lineweaver-Burk plot analysis.

Table 1. Structural statistics of 20 lowest-energy structures of AcCYS.

A. Restraints	
Intra-residue NOEs ($ i-j =0$)	446
Sequential NOEs ($ i-j =1$)	417
Medium range NOEs ($2 \leq i-j \leq 5$)	213
Long range NOEs ($ i-j > 5$)	553
Total NOEs	1629
Dihedral angle restraints	
ϕ	55
ψ	55
Hydrogen bond restraints	43
B. Ensemble Statistics Analysis	
RMSD values (residues 40–135, Å)	
Backbone atoms	0.56±0.12
Heavy atoms	1.11±0.16
Statistics from Ramachandran plot (%)	
Residues in most favored regions	75.4
Residues in additional allowed regions	23.3
Residues in generously allowed regions	1.3
Residues in disallowed regions	0.0

doi:10.1371/journal.pone.0047865.t001

NMR Spectroscopy and Resonance Assignments of AcCYS_DL and AcCYS_DL when Associated with Papain

Backbone resonance assignments of free AcCYS_DL were obtained by correlating intra- and inter-residue through-bond connectivities of $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, and $^{13}\text{C}'$ in a series of standard 2D and 3D NMR spectra including ^{15}N - ^1H HSQC, HNCOC, HN(CA)CO, HNCA, HN(CO)CA, CBCANH, and CBCA(CO)NH. Assignments of H^α , H^β were obtained by ^{15}N -TOCSY-HSQC, HBHA(CACBCO)NH, and HCCH-TOCSY spectra. Resonance assignments of ^{15}N , ^{13}C -labeled AcCYS_DL when associated with unlabeled papain were first performed by comparing chemical shifts with those in free AcCYS_DL. Ambiguous chemical shifts were then obtained by correlating intra- and inter-residue resonances in standard 3D NMR spectra using similar procedures as in the assignments for free AcCYS_DL. Chemical shifts were confirmed by inspection of intra-residue and sequential NOEs in ^{13}C , ^{15}N -edited NOESY-HSQC spectra. Chemical shift perturbation map was used to identify the papain-binding site of AcCYS_DL. A single quantity used as normalized chemical shift perturbation of each residue ($\Delta\delta_{\text{residue}}$) was expressed as equation 3.

$$\Delta\delta_{\text{residue}} = \left\{ \frac{1}{6} \left[(\Delta\delta_{\text{HN}})^2 + (\Delta\delta_{\text{H}^\alpha})^2 + \left(\frac{\Delta\delta_{\text{N}}}{10}\right)^2 + \left(\frac{\Delta\delta_{\text{C}^\alpha}}{4}\right)^2 + \left(\frac{\Delta\delta_{\text{C}^\beta}}{4}\right)^2 + \left(\frac{\Delta\delta_{\text{C}'}}{4}\right)^2 \right] \right\}^{\frac{1}{2}} \quad (3)$$

$\Delta\delta_{\text{HN}}$, $\Delta\delta_{\text{H}^\alpha}$, $\Delta\delta_{\text{N}}$, $\Delta\delta_{\text{C}^\alpha}$, $\Delta\delta_{\text{C}^\beta}$ and $\Delta\delta_{\text{C}'}$ represent the chemical shift difference of AcCYS_DL upon association with papain for nucleus H^N , H^α , N , C^α , C^β , and C' respectively. The secondary structure rearrangement of AcCYS_DL upon association with papain was predicted by CSI program.

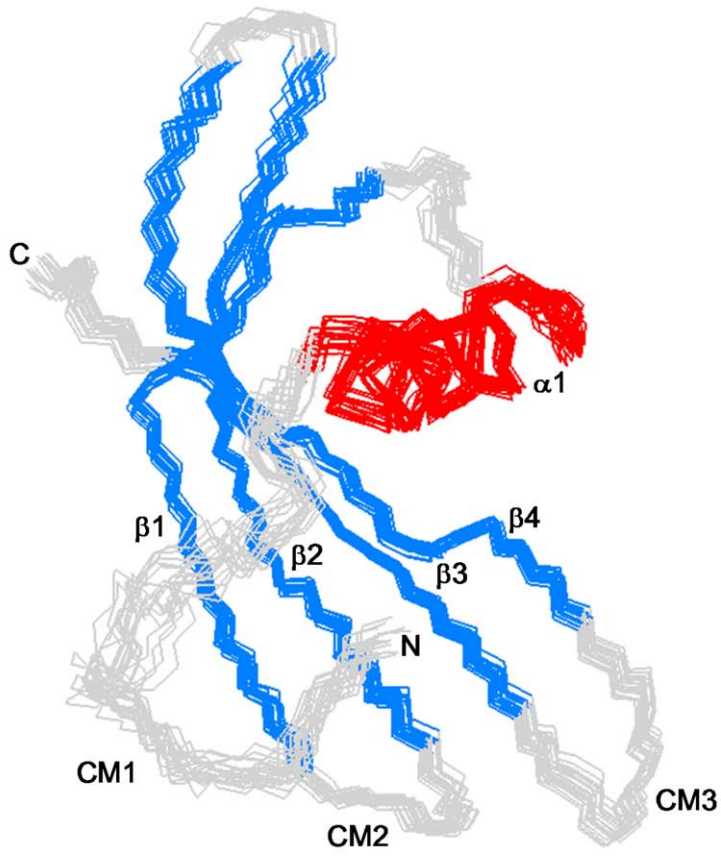
Computational Analysis

Bioinformatics analysis of AcCYS with other cystatin sequences from different families was performed with CLC workbench (CLC bio). Multiple sequence alignments were performed using CLUSTALW algorithm [50]. Prediction of signal peptide was performed using Target P 1.1 server [51]. The solved structure of AcCYS was analyzed and classified by CATH [52] and DALI [53] algorithms.

Molecular Models of AcCYS in Association with Papain

Molecular modeling of AcCYS in association with papain and their docking and restrained molecular dynamic simulation were performed using Discovery Studio 2.1 platform (Accelrys, San Diego, CA). The initial structure model of AcCYS/papain complex was generated from the taro cystatin/papain crystal structure (PDB ID: 3IMA) with the taro cystatin replaced by our NMR AcCYS structure (PDB ID: 2L4V). This complex structure was then energy minimized with CHARMM force field [54]. Docking simulations of AcCYS to papain were performed by use of rigid body ZDOCK 2.1 [55]. The binding site of AcCYS and papain was defined as those residues with significant chemical shifts perturbation upon complex formation and in close contact (≤ 5 Å) with each other in our initial structure model. The orientation of AcCYS was obtained with 6° rotational sampling,

A



B

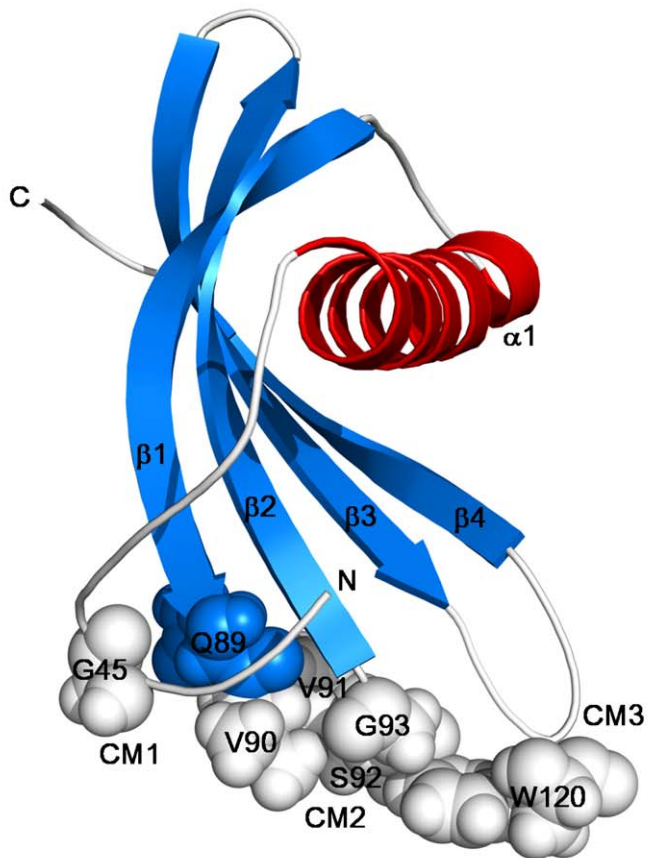


Figure 2. Solution structure of AcCYS. (A) The ensemble of 20 best structures of the inhibitory domain of AcCYS (res 41 to 135) is superimposed. The α -helical, β -sheet, and loop region is color coded in red, green, and grey respectively. (B) Ribbon representation of the inhibitory domain of AcCYS structure shows a $\alpha\beta$ roll structure made up of one α -helix and four anti-parallel β -strands, β_1 – β_4 . Three regions which contain the highly conserved motifs are labeled as CM1–3. Side chains of these highly conserved residues including, G45, Q⁸⁹VVSG, and W120 are shown in spheres. These diagrams were generated using the structural visualization program PYMOL.
doi:10.1371/journal.pone.0047865.g002

and 54000 predictions were generated. The structure with the lowest ZRank score was selected and subsequently subject to energy minimization. The complex structure of AcCYS/papain obtained from ZDock was then further simulated with restrained

molecular dynamics using a cascade protocol in Discovery Studio 2.1. The procedures in the restrained molecular dynamics simulation included minimization with steepest descent and conjugate gradient methods, followed by heating and equilibration

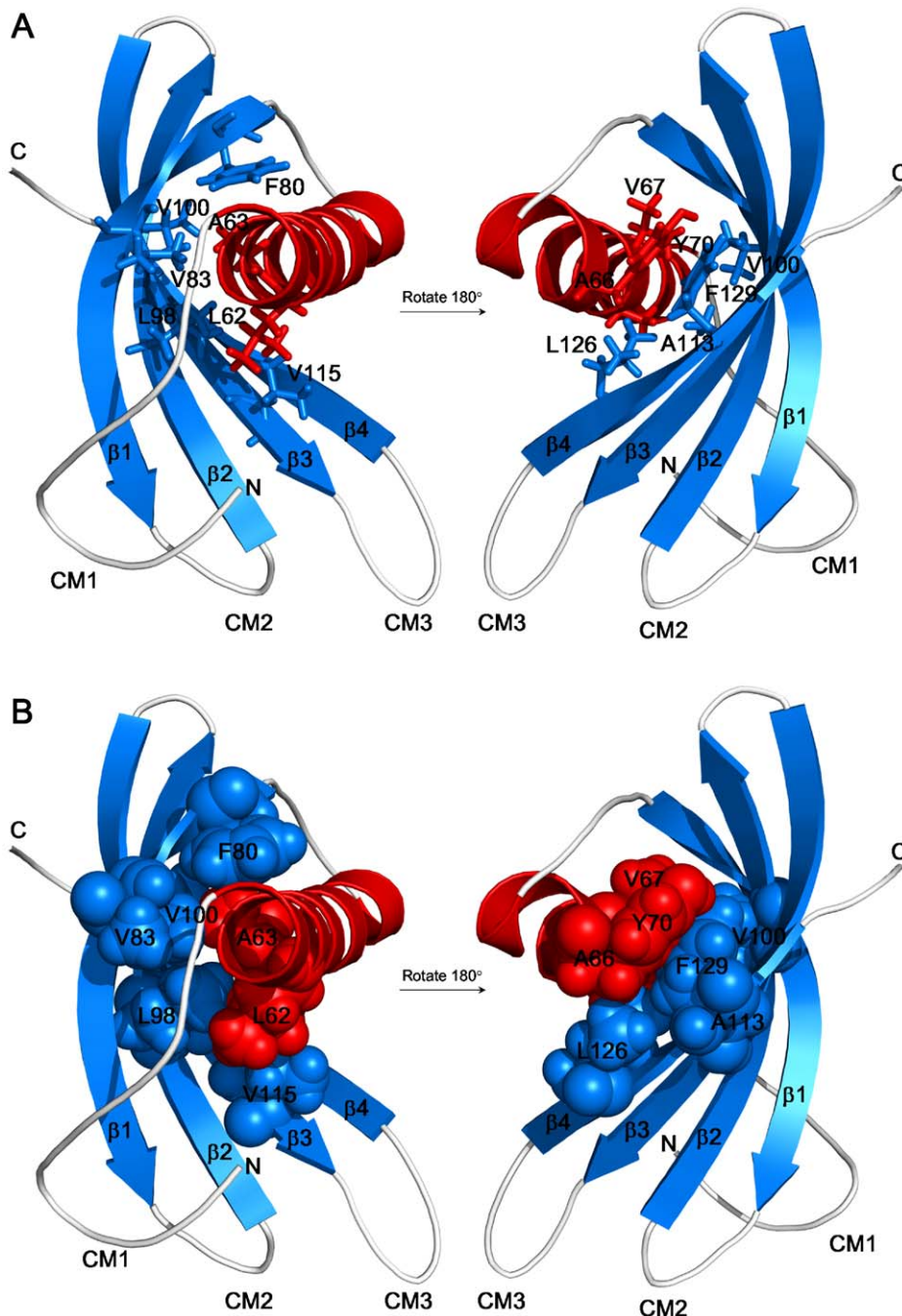


Figure 3. Hydrophobic clusters of AcCYS on the interface of α -helix and β -sheet. Hydrophobic side chains that participate in stabilizing the compact structure of AcCYS are labeled and shown in sticks (A) and in spheres (B).
doi:10.1371/journal.pone.0047865.g003

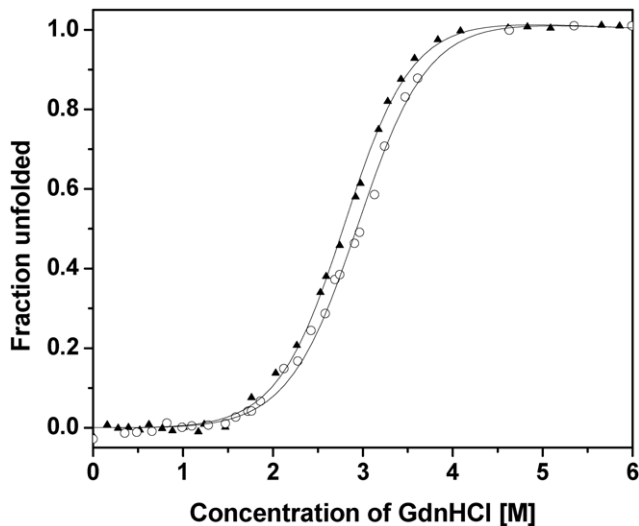


Figure 4. Equilibrium unfolding of AcCYS and AcCYS_DL. The GdnHCl-induced unfolding curves of AcCYS (\blacktriangle) and AcCYS_DL (\circ) were monitored by CD at the wavelength of 222 nm. The transition curve was quantified and the corresponding value of ΔG_D^{H2O} for AcCYS and AcCYS_DL is 4.4, and 4.5 Kcal/mol, respectively. doi:10.1371/journal.pone.0047865.g004

dynamics under 300 K, and finished with production dynamics. The distance restraints used in the molecular dynamics simulation include hydrogen bonds derived from H/D hydrogen exchange and CSI data, intramolecular restraints of hydrophobic clusters in AcCYS, and intermolecular restraints derived from the dock model. The distance-dependent dielectric constant of the solvent was used in the molecular dynamics simulation. The lowest dielectric constant was 1 and the solvent dielectric constant was 80. The conformation of the AcCYS/papain complex with the lowest potential energy in the production dynamics was selected to present the complex structure of AcCYS/papain.

Accession Numbers

^1H , ^{13}C , and ^{15}N chemical shift assignments were deposited in the Biological Magnetic Resonance Data Bank (<http://www.bmrb.wisc.edu>) under accession code 17258. The coordinates of the final ensemble of 20 structures and the NMR restraints used for structural determination were deposited in the PDB under accession code 2L4V.

Results and Discussion

Sequence Alignments

The sequence of AcCYS was aligned with other selective cystatin homologs in Figure 1. Three characteristic motifs of cystatins which were believed to form three-point interaction with target cysteine protease were identified. They included a region containing one or two G residue at the N-terminal region, a QxVxG motif, and a region containing a W residue near C-terminal region. The unique motif LARFAVxExN of phycocystatins was also found in AcCYS. Furthermore, AcCYS possesses an unusual AE-rich sequence of approximately 30 residues in its N-terminus. This AE-rich sequence is unique to AcCYS as verified by extensive blast in databanks.

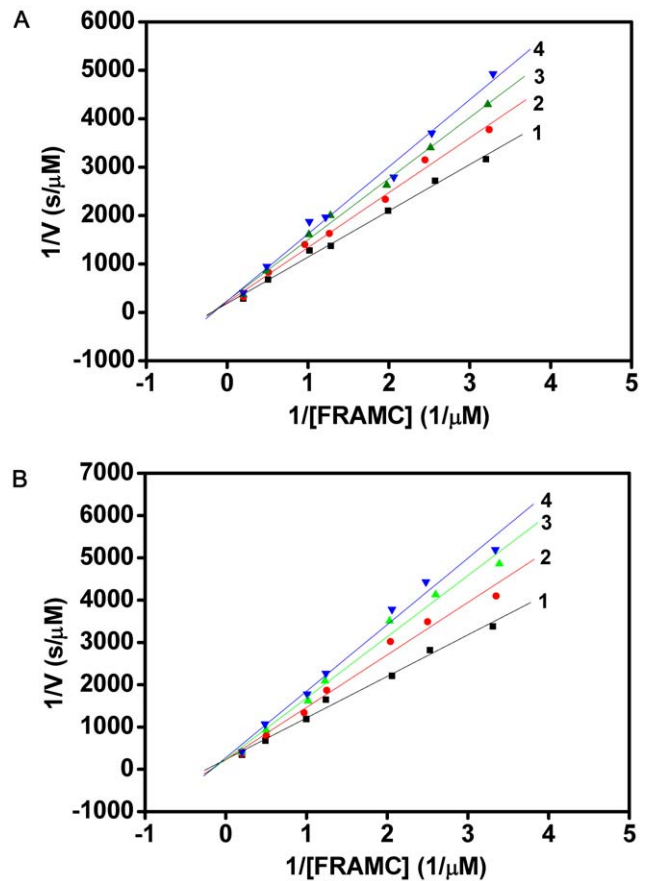


Figure 5. The Lineweaver-Burk plots for the inhibition of papain by AcCYS and AcCYS-DL. The inhibition of papain by AcCYS and AcCYS-DL is shown in (A) and (B) respectively. The initial rates of cleavage of a fluorogenic substrate, Z-Phe-Arg-7-amido-4-methylcoumarin (FRAMC) hydrochloride by papain were obtained spectrofluorometrically with excitation and emission wavelengths at 346 and 450 nm, respectively. Line 1 represents papain activity (2 nM) in the absence of inhibitor. Line 2–4 show the enzyme activity in the presence of 40, 60, 80 pM cystatin, respectively. Both AcCYS and AcCYS-DL strongly inhibited papain with inhibitory constant K_i of $2.0 \pm 0.2 \times 10^{-10}$ M and $1.4 \pm 0.1 \times 10^{-10}$ M, respectively. doi:10.1371/journal.pone.0047865.g005

Structure of Free AcCYS

The solution structures of AcCYS were calculated using CYANA 2.1. There are 568, 913, and 148 NOE distance restraints deduced from the ^{15}N -separated, ^{13}C -separated NOESY-HSQC, and 2D-NOESY for aromatic residues, respectively. In addition to NOE distance restraints, 43 distance restraints for hydrogen bonds and 55 backbone ϕ and ψ restraints were used for final calculation of the structures. An ensemble of 20 conformers with lowest restraints violations was selected for further analysis. Statistical parameters for the 20 calculated structures were listed in Table 1. Root mean square deviations of these structures were 0.56 ± 0.12 Å for backbone atoms and 1.11 ± 0.16 Å for all atoms (residues 40 to 135). The quality of the structural ensemble was evaluated by the program PROCHECK-NMR, in which 98.7% of the residues were in the most favored and additionally allowed regions, only 1.3% in generously allowed region, and no residues in disallowed region in the Ramachandran plot. The final refined ensembles of 20 lowest energy structures have been deposited into the protein data bank (PDB) with accession code 2L4V.

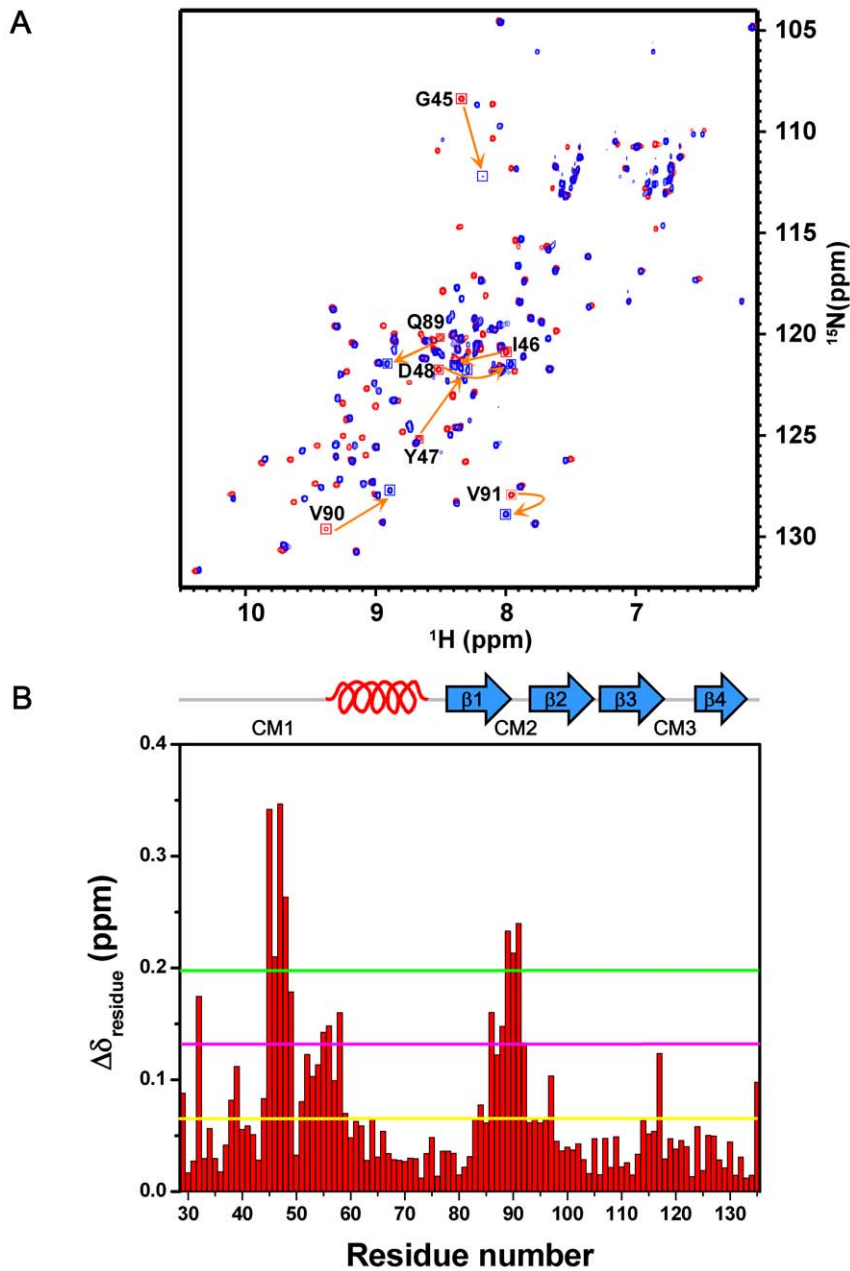


Figure 6. Complex formation of AcCYS_DL and papain. (A) ^{15}N - ^1H HSQC spectra of AcCYS-DL alone (red) and in association with papain (blue). The residues with most significant chemical shift perturbation are labeled. (B) The overall chemical shift perturbation, $\Delta\delta_{\text{residue}}$ of AcCYS_DL upon association with papain was plotted against residue number. $\Delta\delta_{\text{residue}}$ of each residue was calculated as described in Materials and Methods. The average of $\Delta\delta_{\text{residue}}$ for all residues of AcCYS_DL upon association with papain was calculated and shown as the yellow horizontal solid line (0.066 ppm). The pink and green horizontal solid lines represented two and three fold averaged $\Delta\delta_{\text{residue}}$, respectively. doi:10.1371/journal.pone.0047865.g006

Superposition of the 20 best structures of AcCYS together with ribbon diagrams of the averaged structure was shown in Figure 2. AcCYS contains an unstructured N-terminal domain (residues 1 to 40), a loop (residues 41 to 46) and a compact inhibitory domain (residues 47 to 135). The N-terminal AE-rich sequence is located in the unstructured region. The compact inhibitory domain is classified by CATH as a $\alpha\beta$ roll structure comprising an α -helix and a four-stranded β -sheet with the secondary structural topological arrangement of α - β_1 - β_2 - β_3 - β_4 . The α -helix, residues 56–73, is wrapped around by anti-parallel β -strands comprised of residues 79–89 (β_1), 94–104 (β_2), 107–117 (β_3), and

124–131 (β_4). The inner two β -strands, β_2 and β_3 , are coiled smoothly, but the outer two β -strands, β_1 and β_4 , are twisted with three β -bulges on residues V81, V84, and Q127. Together, these bulges produce a tight coiling of the β -sheet, allowing it to wrap around the helix. By DALI 3D-clustering algorithm, the inhibitory domain of AcCYS shows high structural similarity to potato (PDB entry 2W9Q) and taro (PDB entry 3IMA) cystatins with RMSD of 1.9 and 2.1 Å, respectively. However, the short N-terminal β -strand found in residues 24–25 of potato cystatin and residues 6–7 of taro cystatin is not observed in the corresponding residues 46–47 in AcCYS. Three structural motifs (G45, Q⁸⁹XVXG, and

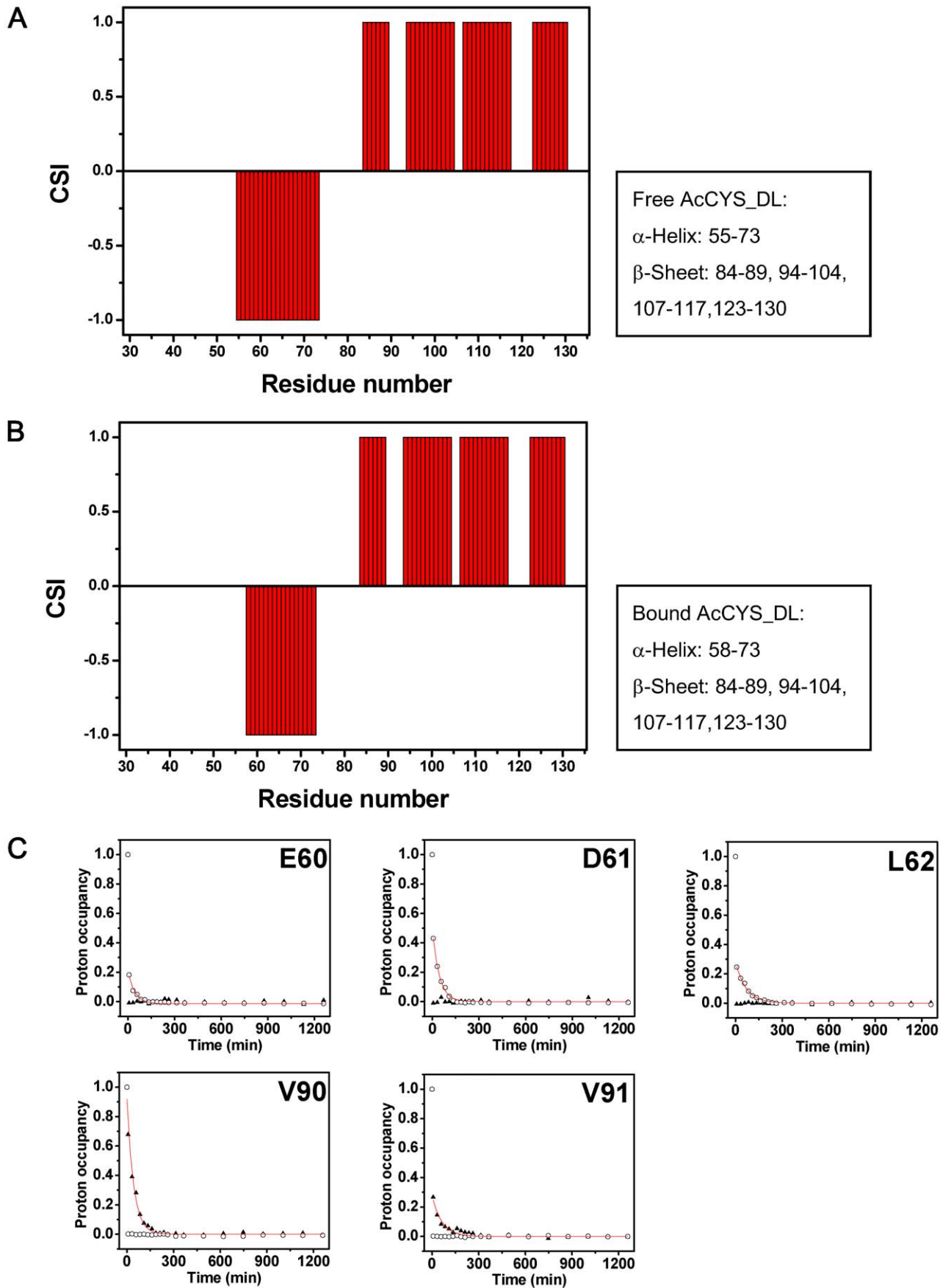


Figure 7. Local secondary structural rearrangements of AcCYS when associated with papain. (A) CSI plot of free AcCYS_DL. (B) CSI plot of AcCYS_DL upon association with papain. (C) H/D exchange of amide protons of V90, V91, E60, D61, and L62 in the bound form (▲) and in the free state of AcCYS DL (○). doi:10.1371/journal.pone.0047865.g007

W120) putatively responsible for the interaction with papain-like proteases are located on one side of AcCYS.

The compact globular structure of AcCYS is mainly stabilized by the hydrophobic clusters on the interface between the α -helix and the β -sheet as shown in Figure 3. The key hydrophobic residues of α -helix, including L62, A63, A66, V67, and Y70, are located on the phycocystatin-specific motif, LARFAVxExN. The structurally complementary hydrophobic residues on β -sheet, including F80, V83, L98, V100, A113, V115, L126, and F129, also show high homology to other phycocystatins.

Secondary Structure and Bulk Thermodynamic Stability Measurements by Circular Dichroism

To assess the effects of the N-terminal extension sequence on the native structure of AcCYS, CD was used to estimate the secondary structure contents of AcCYS and its truncated form, AcCYS_DL (residue 28–135) in the far-UV region. The estimated α -helix and β -sheet contents are 14% and 31% for AcCYS, and 15% and 33% for AcCYS_DL, suggesting that the overall secondary structures of AcCYS are not affected by the removal of the N-terminal AE-rich sequence (residue 1–27). We then compared the stabilities of AcCYS and AcCYS_DL by analyzing their unfolding against denaturant, as monitored by the decrease in the secondary structure content of the proteins. The fraction of unfolded state was plotted as a function of GdnHCl concentration (Figure 4). The unfolding curves of both AcCYS and AcCYS_DL showed cooperative characteristics. The transition curve was quantified by the methods described in Materials and Methods, giving $\Delta G_D^{H_2O}$ values of 4.4, and 4.5 Kcal/mol for AcCYS and AcCYS_DL, respectively. These observations indicate that the removal of N-terminal AE-rich sequence does not affect the thermodynamic stability of AcCYS against GdnHCl-unfolding.

Inhibitory Activity of AcCYS and AcCYS-DL to Papain

A proteolytic activity assay using fluorogenic substrate FRAMC was established to investigate the effect of the N-terminal sequence in AcCYS on its inhibitory activity toward papain. The Lineweaver-Burk plots of the assays with various concentrations of inhibitors AcCYS and AcCYS-DL were shown in Figure 5A and B. The inhibition constants (K_i) for AcCYS and AcCYS-DL to papain were determined according to the plots. Both AcCYS and AcCYS_DL strongly inhibited papain with K_i of $2.0 \pm 0.2 \times 10^{-10}$ M and $1.4 \pm 0.1 \times 10^{-10}$ M, respectively. The results indicate that the removal of the N-terminal sequence in AcCYS has slightly increased the inhibitory ability against papain.

Recognition of AcCYS to Papain Determined by Chemical Shift Perturbation, CSI, and H/D Exchange Studies

Since the unstructured N-terminal region was readily fragmented by papain within NMR acquisition time, we therefore used AcCYS_DL to characterize the interaction. As chemical shifts are sensitive to changes in chemical environments caused either by proximity to the interaction surface or by structural readjustments, they are frequently used as atomic resolution probes to monitor the interaction between biomolecules. We herein used chemical shift perturbation to map the interaction site of AcCYS_DL to papain. Complexes of labeled AcCYS_DL with unlabeled papain were prepared and their NMR spectra were acquired. Backbone

resonance assignments of AcCYS_DL in free and bound forms are completed using the procedures described in Materials and Methods and listed in (Figure S1 and Table S1). The ^{15}N -HSQC of the complex of labeled AcCYS_DL with unlabeled papain was overlaid with the spectrum of free AcCYS_DL (Figure 6A). The majority of the crosspeaks of AcCYS_DL shifted upon association with papain. The normalized chemical shift deviation of each residue, $\Delta\delta_{\text{residue}}$ was calculated as described in Materials and Methods and plotted against residue number as shown in Figure 6B. The average $\Delta\delta_{\text{residue}}$ for all residues of AcCYS_DL upon association with papain was 0.066 ppm. The most perturbed residues of AcCYS_DL of which the $\Delta\delta_{\text{residue}}$ are larger than three folds of the averaged $\Delta\delta_{\text{residue}}$ were found in two conserved loop regions, residue 45 to 48 (GIYD) and residue 89 to 91 (QVV). These loops are likely to be in the proximity to the interaction surface and are predicted to play pivotal roles in the recognition to papain. However, the highly conserved W120 was found to have a chemical shift perturbation less than the average of $\Delta\delta_{\text{residue}}$ and therefore was likely not crucial in the binding to papain. The other strongly perturbed residues of AcCYS_DL with $\Delta\delta_{\text{residue}}$ within two to three folds of averaged $\Delta\delta_{\text{residue}}$ were found in the first turn of α -helical region (residues 55, 56 and 58) and the first strand of the β -sheet (residues 86 and 88). These residues are not located on the primary interaction surface and therefore are suspected to undergo induced fit upon binding of AcCYS_DL to papain. In order to prove the hypothesis, we used chemical shift index (CSI) to detect changes in secondary structure of AcCYS_DL upon binding to papain. The CSI plots of both AcCYS_DL in the free and bound forms showed that the length of the α -helix of AcCYS_DL is shortened upon association with papain while the β -sheet structure remains unchanged (Figure 7A and 7B). In accord with the chemical shift perturbation results, we confirmed that the first turn of the α -helix (residue 56–58) of AcCYS_DL collapses upon association with papain. However, the role of the residues 86 and 88 to the interaction of AcCYS and papain remains unclear.

H/D exchange experiments have been widely applied to map the solvent accessible surface of proteins. The amide protons on the solvent accessible surface of proteins are labile to exchange with solvent deuterium. On the other hand, the most protected amide protons are located in the interior of protein or involved in hydrogen bonds of secondary structure. Hence we used H/D exchange experiments to map the interaction surface and secondary structure changes of AcCYS/papain complex. As shown in Figure 7C, the H/D exchange rates of amide protons of V90, and V91 of AcCYS_DL in the bound form were slower than those in the free form. The amide protons of V90 and V91 on the conserved loop were exposed to solvent in the free form, but protected from exchange with solvent once binding to papain. On the contrary, the H/D exchange rates of amide protons of E60, D61, and L62 of bound AcCYS_DL were much faster than those in the free form. Based on our NMR structure of free AcCYS, the amide protons of residue 60–62 are hydrogen bonded to their respective carbonyl acceptors of residue 56–58 in the helix and therefore protected from exchange with solvent. Rapid exchange of these amide protons with deuterium in the bound form indicated that the hydrogen bonds in the first turn of α -helix break up upon association with papain, further supporting the notion of partial helix collapse. The exchange rates of amide protons of

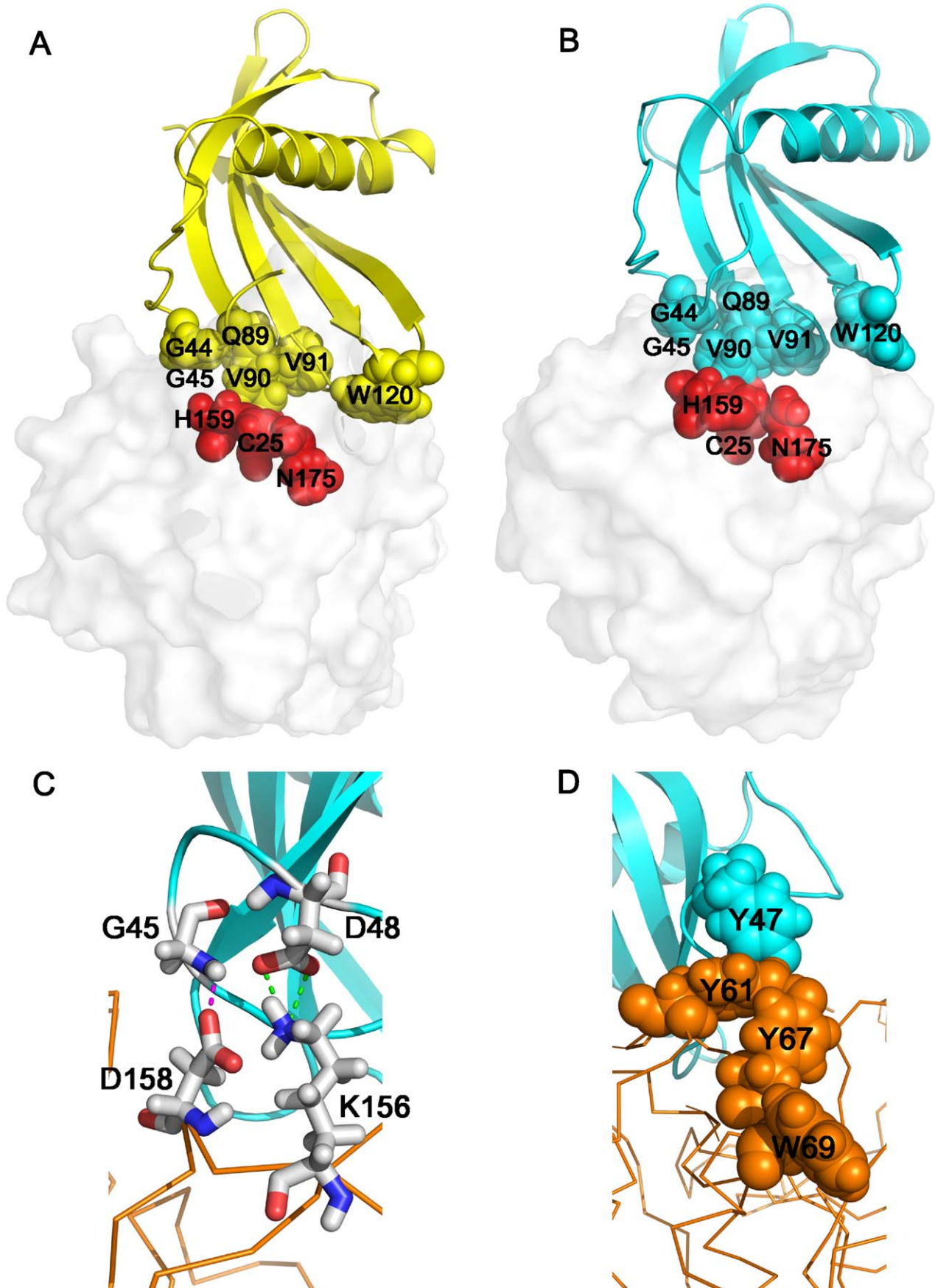


Figure 8. Structural model of AcCYS/papain complex. (A) The complex structure of AcCYS/papain was simulated by ZDOCK. AcCYS is shown in yellow ribbons. The binding site of AcCYS is shown in yellow spheres. The surface of papain is shown in light grey. The active site of papain including C25, H159, and N175 residues are shown in red spheres. (B) Complex structure of AcCYS/papain determined by restrained molecular dynamics simulations. AcCYS is shown in light blue ribbons. The G44–G45 and Q89–V91 of AcCYS shown in spheres completely block the active site of papain. The first turn of α -helix, residue 56–58 collapses. (C) AcCYS/papain complex stabilized by intermolecular forces. The backbone of papain is shown in yellow line. An intermolecular H-bond between G45 amide proton of AcCYS and the side chain carboxyl of D158 of papain and an intermolecular salt bridge between the carboxyl of D48 of AcCYS and K156 of papain are formed in the simulated complex structure (shown in sticks). (D) A new hydrophobic cluster between AcCYS and papain. The side chain of Y47 of AcCYS (shown in light blue spheres) as well as Y61, Y66, and W69 residues of papain (shown in yellow spheres) form an intermolecular aromatic cluster.
doi:10.1371/journal.pone.0047865.g008

residue 45 to 48 (GIYD) were too fast, hence unsuitable as probes in H/D experiments.

In combining the chemical shift perturbation, CSI, and H/D exchange studies, we proposed that the two conserved loop regions of AcCYS, including residues 45 to 48 (GIYD) and residues 89 to 91 (QVV) play primary roles in the recognition to papain, whereas the conserved W120 is not involved. The conformational rearrangement in the helical region of AcCYS is induced upon binding to papain.

Molecular Model for the Interaction between AcCYS and Papain

In order to provide the structural insight for the recognition and inhibition of AcCYS to papain, structures of AcCYS/papain complex were simulated *in silico*. The initial structure model of AcCYS/papain complex was built from the crystal structure of taro cystatin/papain complex with the taro cystatin replaced by our NMR AcCYS structure. Rigid body docking simulation, ZDOCK was used to model the structure of AcCYS/papain complex. The structure with the best ZRank score was selected for further energy minimization. The energy-minimized dock structure of AcCYS/papain is shown in Figure 8A. The residues in the conserved motif 1 and 2 of AcCYS are in proximity to the active site of papain, whereas residue W120 of AcCYS in the third conserved motif protruded from the binding surface. The dock structure of AcCYS/papain was then simulated with restrained molecular dynamics simulation. The restraints include the hydrogen bonds derived from hydrogen exchange and CSI data, intramolecular restraints of hydrophobic clusters in AcCYS, and two intermolecular restraints derived from the dock structure (Table S2). Restrained molecular dynamics simulation was performed for 1 ns and the final simulated AcCYS/papain structure is shown in Figure 8B. The residues in the conserved motif 1 (G44–G45) and 2 (Q89–V91) of AcCYS completely blocked the active site of papain, including C25, H159, and N175. The first turn of the α -helix collapsed. An intermolecular H-bond was found between G45 amide proton of AcCYS and the side chain carboxyl of D158 of papain in the simulated complex structure; the carboxyl of D48 of AcCYS formed a salt bridge with K156 of papain (Figure 8C). In addition, the side chain of Y47 of AcCYS was found to be close to Y61, Y66, and W69 residues of papain and therefore may form a stable intermolecular aromatic cluster to stabilize the complex structure (Figure 8D). These data also explain why the chemical shifts of Y47 and D48 were significantly perturbed upon binding to papain.

Based on the data collected, we propose a molecular model to depict the events following the interaction between AcCYS and papain. Residues V90 and V91 in the second conserved loop (CM2) of AcCYS are responsible for initial contact with papain. Upon engagement, G44 and G45 of the first conserved motif (CM1) move toward the catalytic cleft of papain. This movement pulls down the α -helix, causing it to undergo partial unwinding. Intermolecular forces including salt bridge, H-bond and aromatic

cluster between AcCYS and papain form and stabilize the complex structure. The finely tuned structural rearrangements allow AcCYS to tightly bind to papain and completely block the entry to its active sites.

Conclusions

We have solved the solution structure of the pineapple cystatin by multi-nuclei NMR spectroscopy. A combination of techniques including chemical shift perturbation, CD spectroscopy, and computational modeling was employed to identify sites of contact, determine secondary structure rearrangements, and finally build a model for the interaction between pineapple cystatin and papain. Consistent with other homologous cystatins that bind to the target cysteine proteases by 3-point interactions, the conserved glycine in the N-terminal region and the second hairpin loop of pineapple cystatin make extensive contacts with papain. In contrast, the tryptophan residue in the third conserved loop does not appear to participate in the binding interaction. Overall, binding of pineapple cystatin to papain results in slight rearrangements of the secondary structures, including partial melting of the α -helix and subtle movement of the surface loop toward the active site cleft. These structural perturbations are in turn stabilized by a number of intermolecular forces formed between the two molecules. We also found a non-signal peptide N-terminal extension in pineapple cystatin that is distinctive among other plant species. While this unstructured N-terminal peptide does not seem to affect the overall fold and stability, it reduces the inhibition potency of pineapple cystatin by 30%, likely due to interference of interaction between the first conserved loop of the cystatin and the papain.

Supporting Information

Figure S1 2D ^1H - ^{15}N -HSQC spectrum of 1.0 mM uniformly ^{15}N -enriched AcCYS_DL in 20 mM KCl at pH 6.5, 303 K. Assignments of the backbone amide protons and ^{15}N cross peaks are indicated in the figure. The most crowded region is expanded at the top left corner of the figure for clarity purpose. (TIF)

Table S1 Backbone resonance assignments of free and papain bound forms of AcCYS_DL. (PDF)

Table S2 Intra- and inter-molecular restraints used in the MD simulation of AcCYS_DL/papain complex structure. (PDF)

Author Contributions

Conceived and designed the experiments: FLJT CW CLC. Performed the experiments: DI BC TC TL. Analyzed the data: DI CLC. Wrote the paper: CLC.

References

- Turk B, Turk V, Turk D (1997) Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol Chem* 378: 141–150.
- Kordis D, Turk V (2009) Phylogenomic analysis of the cystatin superfamily in eukaryotes and prokaryotes. *BMC Evol Biol* 9: 266.
- Abrahamson M, Alvarez-Fernandez M, Nathanson CM (2003) Cystatins. *Biochem Soc Symp*: 179–199.
- Vray B, Hartmann S, Hoebeke J (2002) Immunomodulatory properties of cystatins. *Cell Mol Life Sci* 59: 1503–1512.
- Barrett AJ (1986) The cystatins: a diverse superfamily of cysteine peptidase inhibitors. *Biomed Biochim Acta* 45: 1363–1374.
- Margis R, Reis EM, Villeret V (1998) Structural and phylogenetic relationships among plant and animal cystatins. *Arch Biochem Biophys* 359: 24–30.
- Turk V, Bode W (1991) The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett* 285: 213–219.
- Tate S, Ushioda T, Utsunomiya-Tate N, Shibuya K, Ohyama Y, et al. (1995) Solution structure of a human cystatin A variant, cystatin A2–98 M65L, by NMR spectroscopy. A possible role of the interactions between the N- and C-termini to maintain the inhibitory active form of cystatin A. *Biochemistry* 34: 14637–14648.
- Bode W, Engh R, Musil D, Thiele U, Huber R, et al. (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: 2593–2599.
- Martin JR, Craven CJ, Jerala R, Kroon-Zitko L, Zerovnik E, et al. (1995) The three-dimensional solution structure of human stefin A. *J Mol Biol* 246: 331–343.
- Jenko S, Dolenc I, Guncar G, Dobersek A, Podobnik M, et al. (2003) Crystal structure of Stefin A in complex with cathepsin H: N-terminal residues of inhibitors can adapt to the active sites of endo- and exopeptidases. *J Mol Biol* 326: 875–885.
- Stubbs MT, Laber B, Bode W, Huber R, Jerala R, et al. (1990) The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J* 9: 1939–1947.
- Benchabane M, Schluter U, Vorster J, Goulet MC, Michaud D (2010) Plant cystatins. *Biochimie* 92: 1657–1666.
- Cornwall GA, Hsia N (2003) A new subgroup of the family 2 cystatins. *Mol Cell Endocrinol* 200: 1–8.
- Renko M, Pozgan U, Majera D, Turk D (2010) Stefin A displaces the occluding loop of cathepsin B only by as much as required to bind to the active site cleft. *FEBS J* 277: 4338–4345.
- Redzynia I, Ljunggren A, Abrahamson M, Mort JS, Krupa JC, et al. (2008) Displacement of the occluding loop by the parasite protein, chagasin, results in efficient inhibition of human cathepsin B. *J Biol Chem* 283: 22815–22825.
- Redzynia I, Ljunggren A, Bujacz A, Abrahamson M, Jaskolski M, et al. (2009) Crystal structure of the parasite inhibitor chagasin in complex with papain allows identification of structural requirements for broad reactivity and specificity determinants for target proteases. *FEBS J* 276: 793–806.
- Rigden DJ, Mosolov VV, Galperin MY (2002) Sequence conservation in the chagasin family suggests a common trend in cysteine proteinase binding by unrelated protein inhibitors. *Protein Sci* 11: 1971–1977.
- Ljunggren A, Redzynia I, Alvarez-Fernandez M, Abrahamson M, Mort JS, et al. (2007) Crystal structure of the parasite protease inhibitor chagasin in complex with a host target cysteine protease. *J Mol Biol* 371: 137–153.
- Santos CC, Sant'anna C, Terres A, Cunha-e-Silva NL, Scharfstein J, et al. (2005) Chagasin, the endogenous cysteine-protease inhibitor of *Trypanosoma cruzi*, modulates parasite differentiation and invasion of mammalian cells. *J Cell Sci* 118: 901–915.
- Arai S, Matsumoto I, Emori Y, Abe K (2002) Plant seed cystatins and their target enzymes of endogenous and exogenous origin. *J Agric Food Chem* 50: 6612–6617.
- Shyu DJ, Chyan CL, Tzen JT, Chou WM (2004) Molecular cloning, expression, and functional characterization of a cystatin from pineapple stem. *Biosci Biotechnol Biochem* 68: 1681–1689.
- Yang AH, Yeh KW (2005) Molecular cloning, recombinant gene expression, and antifungal activity of cystatin from taro (*Colocasia esculenta* cv. Kaosiung no. 1). *Planta* 221: 493–501.
- Nissen MS, Kumar GN, Youn B, Knowles DB, Lam KS, et al. (2009) Characterization of *Solanum tuberosum* multicystatin and its structural comparison with other cystatins. *Plant Cell* 21: 861–875.
- Chu MH, Liu KL, Wu HY, Yeh KW, Cheng YS (2011) Crystal structure of tarocystatin-papain complex: implications for the inhibition property of group-2 phytocystatins. *Planta* 234: 243–254.
- Nagata K, Kudo N, Abe K, Arai S, Tanokura M (2000) Three-dimensional solution structure of oryzacystatin-I, a cysteine proteinase inhibitor of the rice, *Oryza sativa* L. japonica. *Biochemistry* 39: 14753–14760.
- Shyu DJ, Chou WM, Yiu TJ, Lin CP, Tzen JT (2004) Cloning, functional expression, and characterization of cystatin in sesame seed. *J Agric Food Chem* 52: 1350–1356.
- Neuteboom LW, Matsumoto KO, Christopher DA (2009) An extended AE-rich N-terminal trunk in secreted pineapple cystatin enhances inhibition of fruit bromelain and is posttranslationally removed during ripening. *Plant Physiol* 151: 515–527.
- Pernas M, Sanchez-Monge R, Salcedo G (2000) Biotic and abiotic stress can induce cystatin expression in chestnut. *FEBS Lett* 467: 206–210.
- Abe K, Emori Y, Kondo H, Suzuki K, Arai S (1987) Molecular cloning of a cysteine proteinase inhibitor of rice (oryzacystatin). Homology with animal cystatins and transient expression in the ripening process of rice seeds. *J Biol Chem* 262: 16793–16797.
- Callis J (1995) Regulation of Protein Degradation. *Plant Cell* 7: 845–857.
- Solomon M, Belenghi B, Delledonne M, Menachem E, Levine A (1999) The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. *Plant Cell* 11: 431–444.
- Belenghi B, Acconcia F, Trovato M, Perazzolli M, Bocedi A, et al. (2003) AtCYS1, a cystatin from *Arabidopsis thaliana*, suppresses hypersensitive cell death. *Eur J Biochem* 270: 2593–2604.
- Gutierrez-Campos R, Torres-Acosta JA, Saucedo-Arias IJ, Gomez-Lim MA (1999) The use of cysteine proteinase inhibitors to engineer resistance against potyviruses in transgenic tobacco plants. *Nature Biotechnology* 17: 1223–1226.
- Martinez M, Diaz-Mendoza M, Carrillo L, Diaz I (2007) Carboxy terminal extended phytocystatins are bifunctional inhibitors of papain and legumin cysteine proteinases. *FEBS Lett* 581: 2914–2918.
- Girard C, Rivard D, Kiggundu A, Kunert K, Gledlie SC, et al. (2007) A multicomponent, elicitor-inducible cystatin complex in tomato, *Solanum lycopersicum*. *New Phytol* 173: 841–851.
- Kouzuuma Y, Inanaga H, Doi-Kawano K, Yamasaki N, Kimura M (2000) Molecular cloning and functional expression of cDNA encoding the cysteine proteinase inhibitor with three cystatin domains from sunflower seeds. *J Biochem* 128: 161–166.
- Khaznadji E, Collins P, Dalton JP, Bigot Y, Moire N (2005) A new multidomain member of the cystatin superfamily expressed by *Fasciola hepatica*. *Int J Parasitol* 35: 1115–1125.
- Irene D, Chen BJ, Lo SH, Liu TH, Tzen JT, et al. (2012) Resonance assignments and secondary structure of a phytocystatin from *Ananas comosus*. *Biomol NMR Assign* 6: 99–101.
- Herrmann T, Guntert P, Wuthrich K (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *Journal of Molecular Biology* 319: 209–227.
- Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* 13: 289–302.
- Wishart DS, Sykes BD (1994) Chemical shifts as a tool for structure determination. *Methods Enzymol* 239: 363–392.
- Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 160: 65–73.
- Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54: 905–921.
- Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 14: 51–55, 29–32.
- Laskowski RA, Rullmann JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8: 477–486.
- Whitmore L, Wallace BA (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research* 32: W668–W673.
- Whitmore L, Wallace BA (2008) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* 89: 392–400.
- Pace CN (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* 131: 266–280.
- Higgins DG (1994) CLUSTAL V: multiple alignment of DNA and protein sequences. *Methods Mol Biol* 25: 307–318.
- Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* 300: 1005–1016.
- Orengo CA, Michie AD, Jones S, Jones DT, Swindells MB, et al. (1997) CATH—a hierarchical classification of protein domain structures. *Structure* 5: 1093–1108.
- Holm L, Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* 38: W545–549.
- Brooks BR, Brucoleri RE, Olafson BD, States DJ, Swaminathan S, et al. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry* 4: 187–217.
- Chen R, Li L, Weng Z (2003) ZDOCK: an initial-stage protein-docking algorithm. *Proteins* 52: 80–87.