NMR and crystallographic structural studies of the extremely stable monomeric variant of human cystatin C with single amino acid substitution

Martyna Maszota-Zieleniak, Przemyslaw Jurczak, Marta Orlikowska, Igor Zhukov, Dominika Borek, Zbyszek Otwinowski, Piotr Skowron, Zuzanna Pietralik, Maciej Kozak, Aneta Szymanska and Sylwia Rodziewicz-Motowidlo

1 Faculty of Chemistry, University of Gdansk, Gdansk, Poland
2 NanoBioMedical Centre, Adam Mickiewicz University, Poznan, Poland
3 Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland
4 Department of Biophysics and Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX, USA
5 Department of Macromolecular Physics, Adam Mickiewicz University, Poznan, Poland

Keywords
hCC V57G variant; human cystatin C; NMR; protein structure; X-ray

Correspondence
S. Rodziewicz-Motowidlo and A. Szymańska, Department of Biomedical Chemistry, Faculty of Chemistry, University of Gdansk, Wita Stwosza 63, 80-308 Gdansk, Poland
Tel: +48 58 523 50 37 (SR); +48 58 523 50 34 (AS)
E-mails: s.rodziewicz-motowidlo@ug.edu.pl (SR); aneta.szymanska@ug.edu.pl (AS)

Martyna Maszota-Zieleniak and Przemyslaw Jurczak contributed equally to this work
Aneta Szymańska and Sylwia Rodziewicz-Motowidlo jointly supervised this work

(Received 15 May 2019, revised 14 June 2019, accepted 19 July 2019)
doi:10.1111/febs.15010

Human cystatin C (hCC), a member of the superfamily of papain-like cysteine protease inhibitors, is the most widespread cystatin in human body fluids. This small protein, in addition to its physiological function, is involved in various diseases, including cerebral amyloid angiopathy, cerebral hemorrhage, stroke, and dementia. Physiologically active hCC is a monomer. However, all structural studies based on crystallization led to the dimeric structure formed as a result of a three-dimensional exchange of the protein domains (3D domain swapping). The monomeric structure was obtained only for hCC variant V57N and for the protein stabilized by an additional disulfide bridge. With this study, we extend the number of models of monomeric hCC by an additional hCC variant with a single amino acid substitution in the flexible loop L1. The V57G variant was chosen for the X-ray and NMR structural analysis due to its exceptional conformational stability in solution. In this work, we show for the first time the structural and dynamics studies of human cystatin C variant in solution. We were also able to compare these data with the crystal structure of the hCC V57G and with other cystatins. The overall cystatin fold is retained in the solute form. Additionally, structural information concerning the N terminus was obtained during our studies and presented for the first time.

Database
Crystallographic structure: structural data are available in PDB databases under the accession number 6ROA. NMR structure: structural data are available in PDB and BMRB databases under the accession numbers 6RPV and 34399, respectively.

Abbreviations
BMRB, Biological Magnetic Resonance Bank; GFR, glomerular filtration rate; hCC L68Q, variant of hCC with point mutation L68Q; hCC V57D, variant of hCC with point mutation V57D; hCC V57G, variant of hCC with point mutation V57G; hCC V57N, variant of hCC with point mutation V57N; hCC V57P, variant of hCC with point mutation V57P; hCC, human cystatin C; PDB, Protein Data Base; RMSD, root-mean-square deviation.
Introduction

Human cystatin C (hCC) is a small (13 kDa), 120 amino acid protein and a member of the superfamly of papain-like cysteine protease inhibitors. It is the most widespread cystatin in human body fluids. In addition to its physiological function that is regulation of the activity of inter- and intramolecular cysteine proteases of various origin, it is also involved in numerous diseases, including cerebral amyloid angiopathy, cerebral hemorrhage, stroke, and dementia [1]. The level of HCC is a marker of proper glomerular filtration rate (GFR) [2]. In pathological conditions, hCC co-accumulates with the amyloid β (Aβ) peptide in a form of amyloid deposits, particularly in elderly individuals suffering from Alzheimer’s disease or Down’s syndrome [3]. Furthermore, the leucine 68 to glutamine mutant of cystatin C – L68Q, exhibits high amyloidogenic properties and forms amyloid deposits spontaneously in brain arteries of young adults. This process causes brain hemorrhages and eventually death. The pathological state associated with L68Q hCC accumulation is called Hereditary Cystatin C Amyloid Angiopathy (HCCAA) [3–5].

Under physiological conditions, hCC occurs predominantly as a monomer and only as such is biologically active. In vivo, during cellular trafficking [6], and in vitro at elevated temperatures, low pH or in the presence of low to moderate concentrations of denaturing agents the protein undergoes dimerization [7] and further oligomerization [8]. Both processes are the result of the three-dimensional exchange of the protein domains (3D domain swapping [9]), proceeding in either closed (dimer) or in open-ended (‘run-away’) manner [8,10].

The research on the mechanism of the oligomerization of hCC shows that there are two main factors that contribute to the increased susceptibility of this protein to domain exchange. The first factor is the local strain that is induced by the specific features of the cystatin C amino acid sequence [9,11–13]. The second factor is associated with the presence in the protein structure of a flexible loop L1 which plays the role of a molecular hinge during the domain swapping [14–16]. In order to better understand the dimerization and fibrilization processes of cystatin C (and also other proteins that oligomerize via domain swapping [17–19]), it is very important to know the protein structure and its dynamics in solution. To date, neither the NMR structure nor the dynamic parameters for hCC have been available. Ekiel et al. [20] studied cystatin C with NMR techniques but they mostly focused on the studies of the dimerization process and did not deposit the NMR data. There are, however, the NMR-based structural data available for other members of the cystatin superfamily. The NMR structures for cystatin A (stefin A) and its two mutants – M65L and P25S have been proposed and compared [21,22]. Similar structures for the chicken egg white cystatin [23] and the cystatin from Ananas comosus [24] were also determined. Japelj et al. [25] used NMR techniques to determine the changes in the dynamics of stefin A that occur as a result of 3D domain swapping. The comparison of the above-mentioned structures reveals their high similarity. All the structures contain a cystatin-specific fold – four antiparallel beta-strands forming beta-sheet bent around an alpha-helix. Those structural elements are connected by unordered segments arranged differently in each structure. These differences illustrate a range of the freedom of movements available for the unordered protein segments in aqueous solution.

Such freedom, on the other hand, is limited in the crystal state of a protein. To date, numerous X-ray structures for various members of the cystatin family have been determined. For the wild-type cystatin C only several dimeric structures are available [9,26,27]. Monomeric crystal structures were described for two hCC mutants: V57N (PDB: 3NX0) [15] and stab 1 (PDB: 3GAX) [28]. Both protein variants were also shown to be resistant to dimerization during in vitro experiments exploiting previously established dimerization-promoting conditions [13,14,29]. V57N hCC was one of cystatin C variants designed to study the role of the molecular tensions in the hinge loop L1 and their impact on the stability of the hCC in the monomeric form [12,14]. Variants with aspartic acid and proline in position 57 were also studied. V57D hCC proved to be stable in the desired monomeric form in solution but crystallized as a domain-swapped dimer. V57P hCC was dimeric both in solution and in the crystal lattice, as it was expected [30,31]. In this paper we present the NMR and crystallographic structural data for new hCC hinge loop variant with crucial valine in position 57 changed to glycine. The exceptional stability of the hCC V57G protein allowed us to determine, for the first time, the dynamics of the monomeric hCC in aqueous solution.

Results

Protein expression

The unlabeled proteins were expressed in good yield (up to 5–10 mg of pure protein from 500 mL culture) and purified to homogeneity by applied methods. In the case of isotopically labeled proteins, the obtained
yields were up to 3–5 mg of pure protein from 500 mL of bacterial culture. The yield was much lower especially in the case of the triple labeled proteins. Recent studies by Opitz and coworkers [32] associate the long-observed but poorly studied phenomenon caused by the exchange of H2O to D2O in bacteriological broth with extensive changes in the bacterial (E. coli) proteome leading to both reduction in bacterial growth and influencing protein expression profiles.

Regardless of the labeling type, all proteins could be effectively purified to yield their monomeric forms that was confirmed by the size-exclusion chromatography (Fig. S1).

**Protein stability**

Physiologically, cystatin C is a monomeric protein. However, when exposed to acidic conditions, elevated temperature or in the presence of moderate amounts of chaotropic reagents such as guanidine hydrochloride, it undergoes partial unfolding, followed by domain swapping and formation of a dimer [7]. Also, during the crystallization trials of the wild-type protein, only domain-swapped dimer was obtained, regardless of the condition. The dimerization process of wild-type hCC occurred also during the NMR experiments (performed in our group) and made the obtained data hard to analyze [33]. The efforts to stabilize cystatin C in monomeric, but physiologically active form were undertaken in the Lund group, led by prof. Ander Grubbs, and also in our laboratory. As a result, two structures for the monomeric version of cystatin C were obtained [9,28]. The V57N variant, studied in our group, during the in vitro studies showed low, but still observed, ability to dimerization [14]. In our further efforts, we have focused on finding another hCC variant, preferably with single amino acid substitution, and highly resistant to dimerization. The exchange of the valine in position 57 to glycine provided us with a protein fully fulfilling our requirements. The stability of the hCC V57G variant in the monomeric form at conditions promoting the dimerization of a wild-type cystatin C was verified using gel filtration chromatography. hCC V57G does not exhibit considerable dimerization upon the exposure to temperature up to 60 °C, whereas the wild-type protein dimerizes significantly upon incubation at 50 °C (Fig. 1A). After 24 h of incubation at 60 °C, the dimer content in the hCC V57G sample reached 5% but, at the same time, protein degradation was observed. Acidic conditions, shown to strongly promote dimerization and oligomerization of the wild-type hCC, do not generate dimers of hCC V57G (Fig. 1B). Also, in the presence of an increased concentration of the denaturating agent, that is 0.5 M and 1.0 M guanidine hydrochloride, no significant dimerization in the hCC V57G samples was observed in contrast to the wild-type protein (Fig. 1C) [34].

**NMR experiment**

Structure of the hCC V57G variant was obtained based on two- and three-dimensional nuclear magnetic resonance spectra registered for the double-labeled 13C, 15N protein. First, based on the 2D 1H-15N HSQC spectra (Fig. 2), the chemical shifts for the hydrogen and nitrogen backbone atoms were found. Next, the analysis of the 3D NMR: 3D HNCO, 3D HN(CO)CA, 3D HNCA and 3D CACB(CO)NH was performed.

The analysis of the NMR spectra allowed for the assignment of chemical shifts for 97 out of 114 expected amino acid residues (BMRB accession number 34399). Chemical shifts values were determined for the majority of amino acid residues in the protein sequence, except the most flexible segments – the N terminus and segments of AS loop. Sequence analysis was performed based on chemical shifts for the 13C, 13Cn−1, 13CO, and 13CO4. Next, the NOE analysis of NMR spectra was performed. It required the analysis of the NOESY spectra in the context of the obtained earlier chemical shifts for the main chain atoms of the hCC V57G. As a result, the interprotonal spacing for individual pairs of atoms was determined. The NOE effects of αH-NH(i,−3), αH-βH(i,−3), and αH-NH(i,−4) showed a regular α-helical structure in the Glu21 – Tyr34 region. NMR data (TALOS, NOE) suggest that amino acids from the residues 82 to 91 are also arranged in an α-helix, but it is based mainly on NOE interactions between the side chains. The NOE cross peaks between the residues in the α-helix and the β-sheet show that the five antiparallel strands in hCC V57G variant form a rolled β-sheet surface, rather than a flat one. It is very interesting that in the NOESY spectrum there are several NOE effects between the N-terminal segment of the protein and the amino acid residues found in the AS loop. This observation suggests that the N terminus, despite the undefined structure, weakly interacts with the whole protein molecule.

**NMR structure**

The NMR structure of the hCC V57G protein was based on the NMR data (chemical shifts and NOE effects). The final NMR structures were afterward minimized using small angle X-ray scattering (SAXS)
experimental data. The ensemble view of the NMR structures is shown in Fig. 3. The secondary structure consists of five antiparallel beta-strands: Met14 – Asp15 (β1), Gln48 – Ile56 (β2), Val60 – Leu68 (β3), Ala95 – Val104 (β4), and Thr109 – Thr116 (β5) and two helical segments: (α1) in Glu21 – Tyr34 and (α2) in Asn82 – His90 regions. Both helices are placed perpendicular to each other. The β-strands of the protein are connected with two short loops: L1 (Gly57–Gly59) connecting β-strands β2 and β3, and L2 (Pro105–Gly108) connecting strands β4 and β5. These two loops form a major part of the enzyme binding site [35]. On the opposite side of the molecule, partially disordered structure (appending structure, AS region) connects strands β3 and β4. N and C termini of the protein are disordered. This is particularly evident in the case of the N terminus (see Fig. 3) which is very flexible and tends to form a bent structure at the top of the Gly11-Gly12 residues and the neighboring Pro13. The values of φ and ψ angles for the individual amino acid residues are in the regions of minimum energy (see Table 1) characteristic for the torsion angles for the helical and beta structures. The final NMR structure of the monomeric cystatin C variant V57G was deposited in the Protein Data Bank under the accession code 6RPV.

Fig. 1. Impact of selected conditions on the stability of the hCC V57G variant in the monomeric form: (A) increased temperature, (B) lowered pH and increased temperature, and (C) presence of chaotropic reagent. The wild-type protein was used as a reference. Controls refer to the protein samples analyzed immediately after dissolution at given conditions.
Protein dynamics

$^{15}$N relaxation data ($R_1$, $R_2$, and $^{1}H$-$^{15}$N NOE) for 90 $^{15}$N nuclei out of expected 111 (Fig. S2) were determined. The relaxation data for 21 residues (first eight residues form N terminus included) were not collected due to signals overlap, intensive exchange with water, or low signal-to-noise ratio (Fig. S3). The relaxation parameters for residues located in $\alpha$-helix and $\beta$-sheets are characteristic for the residues in structured regions ($0.80 \pm 0.11$). The measured longitudinal and transverse relaxation rates were $R_1 = 1.17 \pm 0.11$ (s$^{-1}$), $R_2 = 14.52 \pm 3.98$ (s$^{-1}$), respectively (Fig. S2). Analysis of $R_2/R_1$ ratio together with high-resolution 3D structure provided the diffusion constants $D_{II} = 2.53 \times 10^{-7}$ (s$^{-1}$) and $D_{\perp} = 1.65 \times 10^{-7}$ (s$^{-1}$), which indicate considerable anisotropy ($D_{II}/D_{\perp} = 0.65$). The axial symmetrical model of rotation diffusion fits well with the experimental data and exhibits rotational correlation time $\tau_R = 1/(2D_{II} + 4D_{\perp}) = 7.45 \pm 0.11$ ns. This value fits the rotation correlation time expected for proteins with molecular mass c.a. 13.5 kDa.

Relaxation parameters identified residues with relatively low $R_2$ and NOE values, indicating the existence of intensive high-frequency motions in ns – ps time frame facilitating increased backbone flexibility. Such features were demonstrated by Ala58 and Gly59 (Fig. S2). These residues, together with Gly57 (point of mutation) form the loop between $\beta_2$ and $\beta_3$ strands. Another structural loop that shows similar dynamic properties contains residues Trp106–Gly108. These residues are close in space to Gly57–Gly59 segment (Fig. 4).

The analysis of relaxation measurements was performed with spectral density mapping approach (Fig. S3). As described previously, the $J(0)$ vs $J(\omega_A)$ dependence with single motion limit identified residues...
with additional $R_{ex}$ motion (Fig. S4). For instance, slow structural motions have been detected for Asn39, His43, Ser44, Asp81, Gln118, and Asp119 located on the external surface of the hCC V57G structure (Fig. 4). Taking into account the residues which are not observed on $^1$H-$^1$5N HSQC spectra due to the intensive exchange with water we conclude the existence of low frequency (ms – l s range) dynamic motions, responsible for structural rearrangement of the AS protein segment (Fig. 4).

### Crystal structure determination

The crystallization experiments were carried out at 293 K using hanging-drop vapor-diffusion method. First trials for hCC V57G were performed using the commercially available screens and initial conditions were further optimized. Crystals of hCC V57G suitable for data collection appeared within 4 weeks of equilibration against 0.2 M Na acetate, 0.1 M Na cacodylate pH 6.5, and 30% (w/v) PEG 8000. The diffraction data were collected to 2.65 Å resolution and are consistent with space group P6$_1$, with the unit cell parameter $a = 75.83$ Å and $c = 98.18$ Å. The diffraction images were indexed, integrated, and scaled using the HKL program package [36]. The data set was 96.8% complete in the 30–2.65 Å resolution range, and 96.2% complete in the highest resolution shell (Table 2). The Matthews coefficient was 3.04 Å$^3$Da$^{-1}$, indicating two molecules per asymmetric unit with 59.6% of a solvent content. The structure was solved by molecular replacement method using MOLREP [37] and a model of hCC V57N (PDB code 3NX0, [15]) as a search probe. The molecular replacement calculations identified two copies in the asymmetric unit. The model of hCC V57G was refined in REFMAC [38] from the CCP4 package [39]. The refinement converged with a final $R$-factor of 17.44 ($R_{free} = 25.24$) for all data (Table 2). The final model was characterized by the root-mean-square deviation from the ideal bond lengths and angles of 0.013 Å and 1.72°, receptively. Excluding glycine and proline residues, the Ramachandran plot had 94.0% residues in the most favored regions and 5.5% residues in the additionally allowed regions (Table 3). The final X-ray structure of the monomeric cystatin C variant V57G was deposited in the Protein Data Bank under the accession code 6ROA.

The X-ray structure of the hCC V57G variant is shown in Fig. 5. It displays the canonical cystatin fold,
The secondary structure of hCC V57G can be summarized as follows:

(N)-β1-α1-β2-L1-β3-(AS)-β4-L2-β5-(C).

Discussion

Human cystatin C is a biologically important protein that not only plays the vital role of the main inhibitor of cysteine proteases in the human body and is involved in different physiological and pathological processes but is also considered as a reliable marker of certain diseases. Therefore, there is a need to fully characterize this protein in its physiologically relevant monomeric form and describe its structure in solution, which is its natural environment. Previous effort to describe soluble hCC structure did not result in the structural model. We have obtained a new hCC variant with single amino acid substitution in the flexible loop region L1. The exchange of the valine in position 57 with glycine residue provided an exceptionally conformationally stable form of hCC for which we undertook the crystallization and NMR measurements. Both NMR and X-ray crystallographic structures of the hCC V57G variant represent the protein in its monomeric state. This confirms the hypothesis that a point mutation in the region of the L1 loop can stabilize the hCC protein in the monomeric state and inhibit its dimerization process [12]. In both structures, all elements of the cystatins fold are preserved (Fig. 6). The content of a well-defined secondary structure is slightly higher in the X-ray structure, that is reflected in the length of the second, third, and fifth β-strand and α-helix in the protein structure. L1 and L2 loops form in exactly the same regions in both structures. In the NMR structure, an additional, short β-sheet and β-sheet plane in the NMR solution structure are shorter. In addition, the β-sheet plane in the NMR structure has less curvature and the α1-helix is two amino acid residues shorter. Small differences in positions of loops L1 and L2 with respect to each other and the rest of the molecule can be also noted (Figs 6, S5). More significant difference observed between the structures is the position of the AS loop connecting two middle β-strands (β3 and β4). In the case of an NMR structure, the AS loop is further away from the rest of the protein in comparison with the X-ray structure (Figs 6, S5). The structures also differ in the solvent accessible surface, which corresponds to the size of the entire molecule. Solvent accessible surface for NMR structure is 7433.1 Å² (calculated without Ser1-Gly11 residues), whereas for X-ray structure is 6836.3 Å². This may result from the possibility of free

---

Table 3. Geometric analysis of the quality of hCC V57G structures.

|                | NMR       | X-ray     |
|----------------|-----------|-----------|
| RMSD from ideal|           |           |
| Bond lengths (Å) | 0.0150    | 0.0132    |
| Bond angles (°)   | 1.359     | 1.715     |
| Ramachandran statistics |   |           |
| Most favored (%)              | 94.8      | 94.0      |
| Additionally favored (%)      | 5.2       | 5.5       |
| PDB code               | 6RPV     | 6ROA      |

---

five-stranded [Met14 – Asp15 (β1), Tyr42 – Ile56 (β2), Val60 – Thr74 (β3), Ala95 – Val104 (β4), and Thr109 – Asp119 (β5)] antiparallel β-sheet gripped around a long α-helix [Glu21 – Lys36 (α1)]. The secondary structure elements are connected through loops L1 (Gly57 – Gly59) and L2 (Pro105 – Gly108). Additionally, between strands β3 and β4, a broad irregular region called ‘appending structure’ (AS) is positioned.

---

Table 2. Data collection and structure refinement statistics.

| Data collection |                |
|----------------|----------------|
| Space group    | P6₁            |
| Cell parameters|                |
| a, b, c (Å)    | a = b = 75.83, c = 98.18 |
|                | 90.0.90.0 120.0 |
| Temperature (K) | 100            |
| Beamline       | APS 19ID       |
| Wavelength (Å)  | 0.97921        |
| Resolution (Å)  | 2.65 (2.70–2.65) |
| Reflections collected | 579 243 |
| Unique reflections | 9095       |
| Rwork (%)       | 12.7 (56.5)    |
| Rfree (%)       | 17.44/25.24    |
| Completeness (%)| 96.8 (96.2)    |
| Redundancy      | 4.7 (4.3)      |

---

Table 2. Data collection and structure refinement statistics.

Refinement

|                |                |
|----------------|----------------|
| Resolution (Å)  | 2.65           |
| Structure solution method | MOLREP    |
| Molecules/AU   | 2              |
| Rmerge (%)     | 17.44/25.24    |
| Rmerge test set count | 433         |
| Number of atoms| 1741           |
| Protein        | 43             |
| Water          | 43             |
| Average B factor (Å²) | 47.82      |

---

* Values in parentheses correspond to the highest resolution shell.

* Rmerge = Σj |hj - Fj|||Σj|Fj|, where h is the intensity of observation j of reflection h. Rmerge is the coefficient of mean square deviation of the reflection h from the weighted mean.

---

M. Maszota-Zieleniak et al. NMR and X-ray structural studies of hCC V57G protein

The FEBS Journal 287 (2020) 361–376 © 2019 Federation of European Biochemical Societies
movement of this part of the protein in the solution, while in the crystal this part of the protein is tightly packed.

The obtained X-ray structure of the hCC V57G variant was compared with other known monomeric X-ray structures of the cystatin C variants (see Fig. S6): 3GAX (crystal structure of the monomer stabilized with an additional disulfide bridge) and 3NX0 (crystal structure of the hCC monomer with a point mutation V57N). The calculated RMSD values for the structures fitted by the C\text{a} were as follows: 0.65 Å for the hCC V57G/3GAX fitting and 0.32 Å for the hCC V57G/3NX0 fitting (Table S1). The structures of compared proteins are very similar in the arrangement of \alpha-helix and \beta-strands and differ slightly in the arrangement of AS loop.

The NMR structure of the hCC V57G was also compared to known crystal structures of monomeric variants of hCC (Fig. S7). The RMSD values (calculated for residues Pro13-Ala120) were as follows: 3.89 Å for hCC V57G/3GAX and 4.54 Å for hCC V57G/3NX0 pairs, respectively (Table S1). The structures of compared proteins are very similar in the arrangement of \alpha-helix and \beta-strands and differ slightly in the arrangement of AS loop.

The NMR structure of the hCC V57G was also compared to known crystal structures of monomeric variants of hCC (Fig. S7). The RMSD values (calculated for residues Pro13-Ala120) were as follows: 3.89 Å for hCC V57G/3GAX and 4.54 Å for hCC V57G/3NX0 pairs, respectively (Table S1). The structures of compared proteins are very similar – they differ mostly in the loop regions on both poles of the molecule. When residues forming AS loop (Gly69-Lys94) have been omitted from the RMSD, calculations values were 2.19 Å/2.14 Å for 3GAX/3NX0, respectively. The L1 and L2 loops in NMR structure of hCC V57G point to the back of the structure; and the short helix in the AS region seems to be positioned closer to the back of the \beta-sheet plane in comparison to previously known monomeric structures. The N-terminal segment of the protein, undefined in the crystal structure, is visible and adopts mainly extended structure. These differences result from the fact that the NMR structure is determined in solution, therefore has greater freedom of mobility and is more relaxed compared to structures obtained by X-ray crystallography.

Since the X-ray structures are static, it is more appropriate to compare the NMR structure obtained for the hCC V57G variant with other known NMR cystatin structures. Therefore, we compared the hCC V57G NMR structure with 1GD4 (cystatin A with P25S mutation), 1A67 (chicken cystatin), and 2L4V (pineapple cystatin). When the structures were superimposed on all C\text{a} atoms, the calculated RMSD values are as follows: 14.63 Å for hCC V57G/1GD4, 4.36 Å for hCC V57G/1A67, and 6.88 Å for hCCV57G/2L4V pairs (Table S1). Those values are slightly different when RMSD values were calculated using secondary structure matching: 3.07 Å for hCC V57G/1GD4, 2.55 Å for hCC V57G/1A67, and 2.85 Å for hCCV57G/2L4V pairs (Table S1). The similarity in the topology of all compared structures can be observed (Fig. S8). The elements of the secondary structure characteristic for cystatins (the \beta-sheet plane and the helical segment above it) are preserved. Nevertheless, the RMSD values for the fitting of individual structures are high and

Fig. 5. Cartoon representation of the X-ray structure of the hCC V57G in different orientations (PDB code: 6ROA). The Gly57 is shown by the space fill atoms.

Fig. 6. Alignment of NMR (PDB code: 6RPV) with X-ray (PDB code: 6ROA) structure of hCC V67G. Color code: NMR – green, X-ray – cyan.
the structures differ from one another. The differences in RMSD values arise from the fact that each protein belongs to a different cystatin subfamily. This causes differences in the amino acid sequence and the length of the proteins. The greatest sequence similarity occurs between hCC V57G variant and the chicken cystatin structure, which is reflected in the lowest RMSD value. Also, the greatest similarity in the arrangement of the \( \alpha \)-helix pattern in relation to the \( \beta \)-sheet plane can be observed for hCC V57G and chicken cystatin structure. Despite the differences, the overall spatial structures of all the compared proteins are very similar. All the secondary structure elements – \( \alpha \)-helix and five \( \beta \) sheets are in a similar arrangement to each other. The cystatin-specific topology is also preserved. Structures differ primarily in the arrangement and length of AS loops and other unordered segments. This is most probably due to the fact that they are determined based on the data obtained from the experiments performed in solution and exhibit high conformational flexibility.

So far theoretical [12] and structural studies [13,14] show that point mutations could modify the structure and properties of the L1 loop in the human cystatin C. Our results showed that the Asn57 mutations in the L1 loop of hCC could stabilize the closed form of hCC, whereas the Asp57 and Pro57 mutation lead to the opening of the hCC structure and then to dimer/oligomer formation. Structural flexibility of these mutants results most likely from the release of conformational stress in the loop that connects the second and third \( \beta \)-strands in hCC [40,41]. The \( \psi \) angle of the highly conserved Val residue present in the VXG motif is present in the unfavored region on the Ramachandran map in the native protein. In the hCC X-ray structure, stabilized by an additional disulfide bridge (stab 1) [28] and containing the native sequence of the L1 loop, the loop is slightly deformed and the valine side chain is directed to the interior of the loop (Fig. 7). After replacement of residue Val57 by the amino acid without side chain, Gly, the L1 loop shows no distortion (Fig. 7) and no steric hindrance that could destabilize the monomeric form occurs in the L1 loop. In both NMR and X-ray structures, the L1 loop forms a similar structure.

The N-terminal portion of the hCC V57G structure is very flexible but long-range NOEs were identified between the N terminus and the AS structure, which suggests that the N-terminal segment may form a loop in the Val10-Gly11-Gly12-Pro13 region. It is worth mentioning that the N-terminal segment of hCC is cleaved by cysteine proteases. In the case of hCC, cleavage takes place after Gly11 residue. The N-terminally truncated cystatin shows significantly decreased affinity toward cysteine proteases, which indicates that the N-terminal segment is necessary for the inhibitory activity [42,43]. It has been observed that the N-terminal segment of cystatin B [44] or tarocystatin (PDB code:

Fig. 7. The structure of the L1 loop of the hCC V57G NMR structure (green) compared with the X-ray structure of the hCC V57G mutant (cyan) and X-ray structures of hCC stabilized in monomeric form (3GAX – gray, 3NX0 – magenta).

Fig. 8. Alignment of NMR structure of hCC V57G (PDB code: 6RPV) with the structure of tarocystatin in the complex with papain (PDB code: 3IMA). Color code: hCC V57G – green, papain – gray, tarocystatin – blue. The Gly11-Gly12 and OCS (cysteine sulfonic acid) are shown by the space fill atoms. The OCS (yellow) in papain shows the position of cysteine, which belongs to the catalytic triad of the cysteine protease.
3IMA, data not published) in the complex with papain is inserted into the active site of papain. In our NMR structure, the N-terminal segment appears as a bundle of various conformations within a relatively narrow range, which can interact with AS structure (confirmed by NOEs). This arrangement of the N-terminal structure in hCC V57G allows the Val10-Pro13 loop to fit into the active site of the enzyme (papain; Fig. 8). The N-terminal segment of hCC is extremely flexible, but, upon binding to the enzyme, it may form a structure in which Gly11-Gly12 residues are in direct contact with the catalytic site of the enzyme.

The derived overall rotational correlation time of V57G variant of hCC is 7.45 ± 0.11 ns. This value fits the rotation correlation time expected for proteins with molecular mass c.a. 13.5 kDa but is higher than for other cystatins. The derived overall rotational correlation times of P25S and wild-type cystatin A [22] were established as 4.4 ± 0.3 ns and 4.6 ± 0.1 ns, respectively. Other studies of cystatin A show that monomeric and domain-swapped cystatin A the overall rotational correlation times were determined to be 4.6 ± 0.1 ns and 9.2 ± 0.2 ns for the monomer and the dimer, respectively [25]. The differences in the overall rotational correlation time between hCC V57G (120 amino acids) and cystatin A (95 amino acids) are due to the difference in the size of these proteins.

In summary, we have determined the solution structure of the hCC variant, stable in the monomeric form, and have compared it with its crystallographic structure. As a result of the substitution of the valine residue by glycine, the unfavorable deformation of the structure within L1 loops was removed. This resulted in the stabilization of the protein in the monomeric form and gave us the opportunity to determine for the first time the structure and dynamics of the hCC in solution.

Materials and methods

Expression of labeled proteins

The DNA of hCC variant V57G was obtained using site-directed mutagenesis as previously described [14]. Plasmid DNA containing hCC gene, ampicillin resistance gene and temperature promoter was transformed to and expressed in E. coli BL21 (DE3) competent cells (Novagen; Sigma Aldrich Inc., Poznań, Poland). The unlabeled protein used in the crystallization experiments was obtained according to the protocol described earlier [14]. For the expression of double (13C/15N) and triple labeled form (13C/15N/2H) of hCC V57G, a modified protocol by Marley and coworkers was used [45]. Briefly, 500 mL of LB broth (Sigma Aldrich Inc., Poznań, Poland) was inoculated with an overnight culture of transformed E. coli and incubated at 32 °C until the optical density (OD) level reached the value of 0.4 (spectrophotometric measurements, λ = 600 nm). Then, the bacteria were sedimented by centrifugation (10 min/4668 g) and resuspended in 500 mL of M9 minimal medium [45] containing labeled 13C-glucose and 15NH4Cl for double labeling. In the case of the expression of the triple-labeled protein expression was carried out using heavy water (D2O) as a solvent. The culture was further incubated until the OD reached a value of 0.6. Then, the temperature was increased to 42 °C to initiate the protein expression and the incubation was continued at 42 °C for 3 h. Next, the expression was terminated (incubation for 15 min 4 °C), the culture was centrifuged (4 °C, 10 min/4668 g) and the bacterial sediment was stored at −80 °C.

Protein isolation and purification

Expressed proteins were isolated from the bacteria using repeated freeze/thaw treatment followed by classic cold osmotic shock protocol [46]. Protein purification was performed using a two-step chromatographic process, according to the protocol elaborated by Szymańska et al. [14]. In the first step, the ion-exchange column was used (Hitrap™ SP FF, 1 mL, GE Healthcare Poland, Warsaw, Poland) and proteins were eluted using linear salt gradient 0–0.5 m NaCl in 20 mM Tris, pH 7.4. Fractions enriched in cystatin C variant were pooled, dialyzed extensively against 20 mM NH4HCO3, and lyophilized. Next, the second chromatographic step that is gel filtration was performed. Proteins were dissolved in 50 mM NH4HCO3 and separated on a size-exclusion column Superdex™ 75 10/300 (GE Healthcare Poland) run in 50 mM NH4HCO3. Fractions containing purified, monomeric protein (according to gel filtration analysis, see below) were collected, lyophilized, and stored as solid at −20 °C. In both chromatographic step, the stability of solvent flow, linearity of a gradient (when necessary), and the spectrophotometric analysis of the eluate were provided by the AKTA Pure chromatographic system.

Protein analysis

The purity and homogeneity of obtained protein samples, as well as the changes in the latter parameter that is protein dimerization or oligomerization, were monitored using gel filtration chromatography on the Superdex™ 75 PC 3.2/30 column (GE Healthcare Poland) run in 50 mM sodium phosphate, pH 7.4, 150 mM NaCl. The elution profile was observed using spectrophotometric measurement at the wavelength of 280 nm.

Dimerization studies

Thermal dimerization

About 1 mg·mL−1 solution of hCC V57G and hCC wt (used as a control) in PBS buffer (0.01 M phosphate, 0.0027 M KCl, 0.137 M NaCl, pH 7.4; Sigma Aldrich Inc.)
NMR and X-ray structural studies of hCC V57G protein

M. Maszota-Zieleniak et al.

containing 1 mM benzamidine hydrochloride as an internal standard was incubated at 37 °C, 42 °C, 50 °C, and 60 °C with constant orbital shaking.

In another experiment, proteins were dissolved in 50 mM sodium acetate (pH 4.0), 150 mM NaCl buffer to 1 mg·mL⁻¹ concentration and were incubated at 37 °C or 42 °C for 24 or 72 h. Samples were analyzed for the presence of dimers or higher oligomers using gel filtration chromatography.

**Chemical dimerization**

V57G and wild-type cystatin C variants were dissolved at 1 mg·mL⁻¹ in PBS buffer (see above) supplemented with guanidine hydrochloride to 0.5 M or 1.0 M concentrations. Samples were incubated at 37 °C without mixing. The dimerization progress was assessed after 24, 48, 96, and 500 h using analytical gel filtration chromatography.

**Nuclear magnetic resonance measurements**

**NMR sample preparation**

The NMR samples were obtained by dissolution of uniformly labeled (¹⁵N- or ¹³C/¹⁵N-) protein in 90%/10% H₂O/D₂O 50 mM phosphate buffer containing 50 mM NaCl to a final concentration of approximately 0.6 mM (8 mg/mL). The uncorrected pH value was stabilized at 7.4.

**NMR spectroscopy**

The NMR magnetic resonance spectra of the hCC V57G variant were registered with Agilent DDR2 800 MHz spectrometer operated at 18.8 T (¹H resonance frequency 799.94 MHz) at 298 K, installed in NanoBioMedical Centre in Poznań. Spectrometer equipped with four channels, Performa IV z-gradient unit, and ¹H/¹³C/¹⁵N probe head with inverse detection is fully suitable for acquiring the multidimensional NMR data. NMR spectra measurements were performed at 25 °C and 30 °C. The assignments of ¹H, ¹³C, and ¹⁵N backbone resonances were extracted from 3D HNCO/HN(CA)CO/H(CO)CA/HNCA/CBCA(CO)NH/HNACB spectra [47]. Side chains assignments were achieved with C(CO)NH/H(CO)NH/HCC-HTOCSY and ¹⁵N- and ¹³C-edited NOESY experiments. All ¹H, ¹³C, and ¹⁵N chemical shifts were referenced in an indirect manner with respect to external sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) using Ξ = 0.251449530 and 0.101329118 ratio for ¹³C and ¹⁵N resonances, respectively [48]. Recorded NMR data were processed by NMRPipe software [49] and analyzed with SPARKY program [50].

**¹⁵N relaxation measurements**

¹⁵N relaxation data were acquired at 298 K on uniformly ¹⁵N-labeled protein sample on Agilent DDR2 800 spectrometer at 18.8 T (¹H resonance frequency 799.844 MHz). A detailed description of the experiment concerning the analysis of ¹⁵N relaxation data can be found in Supporting materials. The pulse sequence was included in BIOPACK software (Agilent Inc., Palo Alto, USA) written on the basis of previously published experiments [51]. The ¹⁵N R₁ relaxation rates were obtained with ten delays: 10, 90, 170, 290, 410, 550, 690, 850, 1010, and 1250 ms. The ¹⁵N R₂ relaxation rates were calculated with nine delays: 10, 30, 50, 70, 90, 110, 130, 170, and 210 ms. The recycling delay in experiments was kept at 3.5 s. The R₁ and R₂ errors were obtained as standard deviations from 200 Monte Carlo simulations [52]. The steady-state ¹H-¹⁵N heteronuclear NOE were obtained with 6 s relaxation delay from two experiments – with and without ¹H saturation. Errors for NOE values were evaluated from signal-to-noise ratios in recorded spectra [53]. All spectra were processed with NMRPipe [49] and analyzed with SPARKY [54] software.

**NMR structure determination and refinement**

The initial structure of the hCC V57G protein was generated in the RC-Rosetta software [55] on the basis of chemical shift values and homology. The generated structure was then used as a starting structure in the CYANA program [56] (version 3.98.5). In these calculations, 1036 distance constraints (211 intraresidue, 333 sequential, 229 medium-range and 263 long-range) obtained from the analysis of 3D ¹⁵N- and ¹³C-edited NOESY spectra were used. Additionally, the 194 restraints for backbone torsion angles (φ and ψ), together with 44 restraints for χ₁ torsion angles, evaluated with the TALOSn software [57], were also implemented. The 114 distance constraints for 57 hydrogen bonds were defined on the basis of geometric criteria and used only on the final stage of structural refinement. Finally, 20 structures, based on the lowest target function were selected. Further evaluated structures were minimized against small angle X-ray scattering (SAXS) experimental data (see details in the next paragraph) using specific protocol included in XPLOR-NIH program (version 2.47) [58]. In the next stage, structures obtained in XPLOR-NIH were placed into the water box and minimized using the YASARA software [59] (version 19.1.27). The ensemble of 20 structures was finally analyzed with WhatIf [60] software and confirmed high quality of the 3D structure of hCC V57G in solution (Table 1).

**Small angle X-ray scattering**

The SAXS data for hCC V57G variant in solution, which we used as additional data set for NMR structure refinement, were collected using XEUSX 2.0 SAXS/WAXS system (XENOCS, France) and a laboratory X-ray source MetalJet (Excillum AB, Sweden). The data were collected in the high-flux mode and the detector (Pilatus 3R 1M) was located 1200 mm from the sample stage, which
resulted in the scattering vector (s) range covered was from 0.08 to 4.7 nm\(^{-1}\). This SAXS data were used as reference data (fully monomeric data set) in our previous experiments on the radiation-induced domain swapping of hCC [61]. However, data used here were collected for the fully monomeric sample and were analyzed for any radiation damages.

**Crystallization**

Initial crystallization trials for hCC V57G were performed using the commercially available screens Classic, PACT and JCSG+ (Qiagen, Germantown, MD, USA). The conditions were further optimized using reagents generated in-house. All trials were carried out by the sitting-drop vapor-diffusion method at 293 K using EasyXtal plates (Qiagen). Crystallization drops were prepared by mixing 0.6 μL protein solution (10 mg mL\(^{-1}\)) with 0.6 μL well solution on a cover slide. The drops were equilibrated against 500 μL well solution. Crystals of hCC V57G used in the diffraction experiment were obtained using a well solution consisting of 0.2 M Na acetate, 0.1 M Na cacodylate pH 6.5 and 30% (w/v) PEG 8000. Before cryocooling crystals were soaked in reservoir solution supplemented with 30% (v/v) PEG 400 for cryoprotection and then plunged into liquid nitrogen.

**X-ray data collection and processing**

X-ray diffraction data were collected at 100 K using synchrotron radiation on beamline 19ID of the Advanced Photon Source (APS), Argonne National Laboratory in Chicago on ADSC Quantum Q315 CCD detector. The data collection statistics are summarized in Table 2.

About 240 frames were collected using oscillation range 0.5° for a crystal-to-detector distance of 319 mm with exposure time was 1 s. A total of 579 243 reflections were measured and reduced to 9,095 unique data extending to 2.65 Å resolution. This data set is 96.8% complete (96.2% in the last resolution shell) and is characterized by \(R_{merge}\) of 12.8 and \(<\Delta I/\sigma I>\) of 13.7. The diffraction images were indexed, integrated, and scaled using the HKL-3000 package [36,62].

**X-ray structure determination and refinement**

The structure was solved by molecular replacement method using MOLREP [37] and monomer of hCC V57N as a search probe (PDB code: 3NX0 [15]). Two protein molecules were found in the asymmetric unit of the P\(_6_1\) unit cell. The model was refined in REFMAC [38] from the CCP4 package [39]. Several cycles of manual model rebuilding in the electron density maps were performed in COOT [63]. Problematic loop regions were corrected and waters were added manually in COOT. The TLS groups were defined according to the TLSMD server [64]. The final model of hCC V57G contained 43 water molecules. The progress of the X-ray structure refinement was monitored and the model was validated using the \(R_{free}\) parameter [65]. The structures refinement statistics are summarized in Table 2.

**Quality of the NMR and X-ray structures**

Programs PROCHECK [66] and the MolProbity server [67] were used to assess the quality of the final NMR and X-ray structure. The Ramachandran statistics and RMSD values are summarized in Table 3. Atomic coordinates and structure factors of the NMR and X-ray structures have been deposited in the RCSB Protein Data Bank under accession numbers 6RPV and 6ROA, respectively. Analysis and visualizations of the obtained 3D structures were carried out using and pymol [68] software. The solvent accessible surface area was calculated with the molsol program [69]. The solvent radius was 1.4 Å and the precision was 3 Å.

**Acknowledgements**

This work is financially supported by the grant from the National Science Centre Poland UMO/2011/03/N/ST4/01293. Diffraction data were collected at Advanced Photon Source (APS), Argonne National Laboratory, Structural Biology Center (SBC) at the beamline 19ID. SBC-CAT is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. The SAXS data refinement was financed by the bilateral research and development grant: POLTUR2/3/2017 from National Centre for Research and Development (Poland) and 117Z009 from TÜBİTAK (Turkey). We thank Izabela Behrendt, Paulinaペット, and Sławomir Lach for the technical support. PJ thanks the “Polish Returns” programme project (No PPN/PPO/2018/1/00071), from the National Agency for Academic Exchange, in which he is currently employed.

**Conflict of interest**

The authors declare no conflict of interest.

**Author contributions**

MM-Z provided financial support, designed and performed NMR studies, analyzed the results, wrote the manuscript. PJ obtained the proteins for NMR and X-ray studies, analyzed the results, wrote the manuscript. MO designed and performed crystallization experiments and X-ray structure determination, analyzed the
results, wrote the manuscript. IZ performed NMR spectra and NMR structure calculations. DB collected and processed X-ray data. ZO collected and processed X-ray data. PS helped to obtain the proteins. ZP: collected and analyzed SAXS data. MK collected and analyzed SAXS data. AS conceived and designed the study, analyzed the results, wrote the manuscript. SR-M provided financial support, created the study concept, designed experiments, supervised NMR studies, analyzed the results, revised the manuscript. All authors read and approved the final manuscript.

References

1 Levy E, Jaskolski M & Grubb A (2006) The role of cystatin C in cerebral amyloid angiopathy and stroke: cell biology and animal models. Brain Pathol 16, 60–70.

2 Grubb AO (2000) Cystatin C – properties and use as diagnostic marker. Adv Clin Chem 35, 63–99.

3 Olafsson I, Grubb A, Olafson I & Grubb A (2000) Hereditary cystatin C amyloid angiopathy. Amyloid 7, 70–79.

4 Jensson O, Gudmundsson G, Arnason A, Blöndal H, Petursdottir I, Thorsteinsson L, Grubb A, Löffberg H, Cohen D & Frangiome B (1987) Hereditary cystatin C (gamma-trace) amyloid angiopathy of the CNS causing cerebral hemorrhage. Acta Neurol Scand 76, 102–114.

5 Palsdottir A, Abrahamson M, Thorsteinsson L, Arnason O, Olafsson I, Grubb A & Jensson O (1989) Mutation in the cystatin C gene causes hereditary brain hemorrhage. Prog Clin Biol Res 317, 241–246.

6 Merz GS, Benedikz E, Schwenk V, Johansen TE, Vogel LK, Rushbrook JI & Wisniewski HM (1997) Human cystatin C forms an inactive dimer during intracellular trafficking in transfected CHO cells. J Cell Physiol 173, 423–432.

7 Ekiel I & Abrahamson M (1996) Folding-related dimerization of human cystatin C. J Biol Chem 271, 1314–1321.

8 Wahlbom M, Wang X, Lindström V, Carlemalm E, Jaskolski M & Grubb A (2007) Fibrillogenic oligomers of human cystatin C are formed by propagated domain swapping. J Biol Chem 282, 18318–18326.

9 Janowski R, Kozak M, Jankowska E, Grzonka Z, Grubb A, Abrahamson M & Jaskolski M (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. Nat Struct Biol 8, 316–320.

10 Östner G, Lindström V, Hjort Christensen P, Kozak M, Abrahamson M & Grubb A (2013) Stabilization, characterization, and selective removal of cystatin C amyloid oligomers. J Biol Chem 288, 16438–16450.

11 Rodziewicz-Motowidło S, Wahlbom M, Wang X, Lagiewka J, Janowski R, Jaskólski M, Grubb A & Grzonka Z (2006) Checking the conformational stability of cystatin C and its L68Q variant by molecular dynamics studies: Why is the L68Q variant amyloidogenic? J Struct Biol 154, 68–78.

12 Rodziewicz-Motowidło S, Iwaszkiewicz J, Sosnowska R, Czaplewskas P, Sobolewski E, Szymańska A, Stachowiak K & Liwo A (2009) The role of the Val57 amino-acid residue in the hinge loop of the human cystatin C. Conformational studies of the beta2-L1-beta3 segments of wild-type human cystatin C and its mutants. Biopolymers 91, 373–383.

13 Szymańska A, Jankowska E, Orlikowska M, Behrendt I, Czaplewskas P & Rodziewicz-Motowidło 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.

14 Szymańska A, Radulska A, Czaplewskas P, Grubb A, Grzonka Z & Rodziewicz-Motowidło S (2009) Governing the monomer-dimer ratio of human cystatin C by single amino acid substitution in the hinge region. Acta Biochim Pol 56, 455–463.

15 Orlikowska M, Jankowska E, Kołodziejezyk R, Jaskolski M & Szymańska A (2011) Hinge-loop mutation can be used to control 3D domain swapping and amyloidogenesis of human cystatin C. J Struct Biol 173, 406–413.

16 Orlikowska E, Jankowska E, Borek D, Kołodziejezyk R, Otwinowski Z, Jaskolski M & Szymańska A (2010) Effect of point mutation in the hinge region on a structure of an amyloidogenic protein - human cystatin C. J Pept Sci 16, 110–111.

17 Bennett MJ, Sawaya MR & Eisenberg D (2006) Deposition diseases and 3D domain swapping. Structure 14, 811–824.

18 Zerovnik E, Stoka V, Mirtic A, Guncar G, Grdadolnik J, Staniforth RA, Turk D & Turk V (2011) Mechanisms of amyloid fibril formation - focus on domain-swapping. FEBS J 278, 2263–2282.

19 van der Wel PCA (2012) Domain swapping and amyloid fibril conformation. Prion 6, 211–216.

20 Ekiel I, Abrahamson M, Fulton DB, Lindahl P, Storer AC, Levadoux W, Lafrance M, Labelle S, Pomerleau Y, Groeleau D et al. (1997) NMR structural studies of human cystatin C dimers and monomers. J Mol Biol 271, 266–277.

21 Tate S, Ushioda T, Utsunomiya-Tate N, Shibuya K, Ohyama Y, Nakano Y, Kaji H, Inagaki F, Samejima T & Kainosho M (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.

22 Shimba N, Kariya E, Tate S, Kaji H & Kainosho M (2000) Structural comparison between wild-type and
NMR and X-ray structural studies of hCC V57G protein

M. Maszota-Zieleniak et al.

P25S human cystatin A by NMR spectroscopy. Does this mutation affect the α-helix conformation? J Struct Genetomics 1, 26–42.

23 Dieckmann T, Mitschang L, Hofmann M, Kos J, Turk V, Auerswald EA, Jaenicke R & Oschkinat H (1993) The structures of native phosphorylated chicken cystatin and of a recombinant unphosphorylated variant in solution. J Mol Biol 234, 1048–1059.

24 Irene D, Chen B-J, Lo S-H, Liu T-H, Tzen JT-C & Chyan C-L (2012) Resonance assignments and secondary structure of a phytocystatin from Ananas comosus. Biomol NMR Assign 6, 99–101.

25 Janowski R, Mitschang L, Hofmann M, Kos J, Turk V, Auerswald EA, Jaenicke R, and Oschkinat H (1993) The structures of native phosphorylated chicken cystatin and of a recombinant unphosphorylated variant in solution. J Mol Biol 234, 1048–1059.

26 Irene D, Chen B-J, Lo S-H, Liu T-H, Tzen JT-C & Chyan C-L (2012) Resonance assignments and secondary structure of a phytocystatin from Ananas comosus. Biomol NMR Assign 6, 99–101.

27 Janowski R, Kozak M, Abrahamson M, Grubb A & Janowski R, Kozak M, Abrahamson M, Grubb A & Jaskolski M (2005) 3D domain-swapped human cystatin C with amyloidoid-like intermolecular beta-sheets. Proteins 61, 570–578.

28 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, 1726–1737.

29 Nilsson M, Wang X, Rodziewicz-Motowidlo S, Janowski R, Lindström V, Onnerfjord P, Westmark G, Grzonka Z, Jaskolski M & Grubb A (2004) Prevention of domain swapping inhibits dimerization and amyloid fibril formation of cystatin C: use of engineered disulfide bridges, antibodies, and carboxymethylpapain to stabilize the monomeric form of cystatin C. J Biol Chem 279, 24236–24245.

30 Orlikowska M, Szmyriska A, Borek D, Otwinowski Z, Skowron P & Jankowska E (2013) Structural characterization of V57D and V57P mutants of human cystatin C, an amyloidogenic protein. Acta Crystallogr D Biol Crystallogr 69, 577–586.

31 Orlikowska M, Jankowska E, Borek D, Otwinowski Z, Skowron P & Jankowska E (2011) Crystallization and preliminary X-ray diffraction analysis of Val57 mutants of the amyloidogenic protein human cystatin C. Acta Crystallogr Sect F Struct Biol Cryst Commun 67, 1608–1611.

32 Opitz C, Ahrné E, Goldie KN, Schmidt A & Grzesiek S (2019) Deuterium induces a distinctive Escherichia coli proteome that correlates with the reduction in growth rate. J Biol Chem 294, 2279–2292.

33 Maszota-Zieleniak M (2017) Struktura i dynamika peptydów i białek amyloidogenicznach na przykładzie serum amyloidu A i ludzkiej cystatyny C. Ph.D. Thesis, Faculty of Chemistry, University of Gdańsk.

34 Behrendt I (2010) Badanie wpływu mutacji Val57-Gly na stabilność konformacją monomeru ludzkiej cystatyny C. M.Sc Thesis, Faculty of Chemistry, University of Gdańsk.

35 Turk V, Stoka V & Turk D (2008) Cystatins: biochemical and structural properties, and medical relevance. Front Biosci 13, 5406–5420.

36 Otwinowska Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 276, 307–326.

37 Vagin A & Teplyakov A (1997) MOLREP: an Automated Program for Molecular Replacement. J Appl Crystallogr 30, 1022–1025.

38 Musshudov GN, Skubák P, Lebedev AA, Panu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC 5 for the refinement of macromolecular crystal structures. Acta Crystallogr Sect D Biol Crystallogr 67, 355–367.

39 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr Sect D Biol Crystallogr 67, 235–242.

40 Engh RA, Dieckmann T, Bode W, Auerswald EA, Turk V, Huber R & Oschkinat H (1993) Conformational Variability of Chicken Cystatin: Comparison of Structures Determined by X-ray Diffraction and NMR Spectroscopy. J Mol Biol 234, 1060–1069.

41 Stanforth RA, Giannini S, Higgins MJ, Conroy MJ, Hounslow AM, Jerala R, Craven CJ & Waltho JP (2001) Three-dimensional domain swapping in the folded and molten-globule states of cystatins, an amyloid-forming structural superfamily. EMBO J 20, 4774–4781.

42 Abrahamson M, Ritonja A, Brown MA, Grubb A, Macleidt W & Barrett AJ (1987) Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. J Biol Chem 262, 9688–9694.

43 Abrahamson M, Mason RW, Hanson H, Buttle DJ, Grubb A & Ohlsson K (1991) Human cystatin C. role of the N-terminal segment in the inhibition of human cysteine proteinases and in its inactivation by leucocyte elastase. Biochem J 273 (Pt 3), 621–626.

44 Stubbs MT, Laber B, Bode W, Huber R, Jerala R, Lenarcic B & Turk V (1990) The refined 2.4 A 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.

45 Bracken C, Marley J & Lu M (2001) A method for efficient isotopic labeling of recombinant proteins. J Biomol NMR 20, 71–75.
Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Superposition of hCC V57G structures (NMR and X-ray) on the existing cystatin models.

Fig. S1. Size-exclusion chromatogram of the purified wild-type cystatin C and its V57G variant hCC proteins in differently labeled versions.

Fig. S2. $^{15}$N relaxation data (R1, R2 and $^{1}H, ^{15}N$ NOE) determined for hCC V57G protein on 18.8 T at 298 K.

Fig. S3. The ribbon representation of 3D structure of the hCC V57G variant.

Fig. S4. Graphical analysis of spectral density values.

Fig. S5. Distance between Cα atom of each residue of the X-ray structure of hCC V57G and equivalent Cα atom of NMR structure of hCC V57G.
**Fig. S6.** The X-ray structure of the hCC V57G mutant (cyan) compared with the known X-ray hCC monomer structures: 3GAX (gray) and 3NX0 (magenta).

**Fig. S7.** The NMR structure of the hCC V57G mutant (green) compared with the known X-ray hCC monomer structures: 3GAX (gray) and 3NX0 (magenta).

**Fig. S8.** The NMR structure of the hCC V57G mutant (green) compared with other known NMR cystatin structures: 1GD4 (blue), 1A67 (yellow) and 2L4V (brown).

**Appendix S1.** Analysis of $^{15}$N relaxation data.