Crystal structure of the SENP1 mutant C603S–SUMO complex reveals the hydrolytic mechanism of SUMO-specific protease

Zheng XU, So Fun CHAU, Kwok Ho LAM, Ho Yin CHAN, Tzi Bun NG and Shannon W. N. AU¹

Department of Biochemistry and Molecular Biotechnology Program, Centre for Protein Science and Crystallography, Faculty of Science, The Chinese University of Hong Kong, Shatin, Hong Kong

SUMO (small ubiquitin-related modifier)-specific proteases catalyse the maturation and de-conjugation processes of the sumoylation pathway and modulate various cellular responses including nuclear metabolism and cell cycle progression. The active-site cysteine residue is conserved among all known SUMO-specific proteases and is not substitutable by serine in the hydrolysis reactions demonstrated previously in yeast. We report here that the catalytic domain of human protease SENP1 (SUMO-specific protease 1) mutant SENP1C^{C603S} carrying a mutation of cysteine to serine at the active site is inactive in maturation and de-conjugation reactions. To further understand the hydrolytic mechanism catalysed by SENP1, we have determined, at 2.8 Å resolution (1 Å = 0.1 nm), the X-ray structure of SENP1C^{C603S}–SUMO-1 complex. A comparison of the structure of SENP2–SUMO-1 sug-

gests strongly that SUMO-specific proteases require a self-conformational change prior to cleavage of peptide or isopeptide bond in the maturation and de-conjugation processes respectively. Moreover, analysis of the interface of SENP1 and SUMO-1 has led to the identification of four unique amino acids in SENP1 that facilitate the binding of SUMO-1. By means of an *in vitro* assay, we further demonstrate a novel function of SENP1 in hydrolysing the thioester linkage in E1-SUMO and E2-SUMO complexes. The results disclose a new mechanism of regulation of the sumoylation pathway by the SUMO-specific proteases.

Key words: conjugation, cysteine protease, de-conjugation, maturation, small ubiquitin-related modifier (SUMO), thioester linkage.

INTRODUCTION

SUMO (small ubiquitin-related modifier) is a member of the ubiquitin-like protein family that regulates the cellular function of a variety of target proteins in all eukaryotic organisms. In contrast with ubiquitination, which contributes to regulation of apoptosis, the biological consequences of SUMO modification include nuclear targeting, formation of subnuclear structures, regulation of transcriptional activities and control of protein stability [1,2]. In recent years, many SUMO substrates have been identified in human and yeast using proteomic strategies [3,4]. These new substrates have furthered our understanding of sumoylation as an important post-translational modification in a diverse array of biological processes. There is increasing evidence to show that sumoylation is able to affect or be affected by other post-translational modifications such as phosphorylation and ubiquitination [5–8].

SUMOs are highly conserved in all species, from yeast to human. In vertebrates, there are three SUMO family members, namely, SUMO-1, -2 and -3, whereas in invertebrates there is only a single SUMO gene [9]. Recently, another SUMO member, SUMO-4, has been identified in humans [10], which has been characterized in the maturation process [11]. SUMO proteins are expressed in their precursor forms and SUMO-specific proteases catalyse the maturation process to expose the C-terminal Gly-Gly motif. SENP1 (SUMO-specific protease 1) and SENP2 manifest distinct maturation efficiencies for different SUMOs [12,13]. Like the ubiquitination system, SUMO is activated by the E1 enzyme (which in yeast is composed of the heterodimer of Uba2 and Aos1), transferred to E2 (Ubc9) and conjugated to its substrates with the help of E3-like enzymes [1,2]. Interestingly, an unusual

adenoviral protein, Gam1, has been demonstrated to interfere with the activity of E1 and abrogate the SUMO conjugation reaction [14]. In a reversible reaction, SUMO proteases are able to de-conjugate SUMO from its substrates to regulate the level of sumoylation [13,15].

To date, seven SUMO specific proteases have been identified in humans, which have been designated SENP1–7. All seven proteases possess a conserved C-terminal domain, although these are variable in size and a distinct N-terminal domain. The C-terminal domain exhibits catalytic activity, while the N-terminal domain regulates cellular localization and substrate specificity [16,17]. The crystal structures of the catalytic domains of SENP2 and the SENP2–SUMO-1 complex from humans and Ulp1–SMT3 complex from yeast have been elucidated [13,18]. These structures showed that the active sites of the SUMO-specific proteases resemble those of other cysteine proteases. The nucleophilic active-site cysteine is co-ordinated by a general base histidine residue, which is in turn stabilized by an aspartate residue [13]. A comparison of SENP2 with the SENP2–SUMO-1 complex revealed that SENP2 undergoes local structural rearrangements in response to SUMO-1 binding [13]. Whether all the structural rearrangements occur before or after cleavage of the peptide bond, to facilitate the hydrolysis reaction or to release SUMO, is still not clear.

To gain a better insight into the molecular basis of maturation and de-conjugation processes catalysed by SENP1, we have determined the X-ray structure of a complex between SENP1C (SENP1 catalytic domain; residues 427–644) mutated at C603S and matured SUMO-1, at 2.8 Å (1 Å = 0.1 nm) resolution. The structure shows that the mutant protease does not possess any hydrolytic activity because the substituted serine residue does not

Abbreviations used: Caps, 3-(cyclohexylamino)propane-1-sulphonic acid; CNS, crystallography and NMR system; DALI, distance matrix alignment; GST, glutathione S-transferase; PEG, poly(ethylene glycol); rmsd, root mean square deviation; SENP1, small ubiquitin-related modifier-specific protease 1; SENP1C, SENP1 catalytic domain; SUMO, small ubiquitin-related modifier; TDG, thymine DNA glycosylase.

¹ To whom correspondence should be addressed (email Shannon-au@cuhk.edu.hk).

undergo any local structural rearrangements at the active site as observed in the SENP2–SUMO-1 complex [13]. This suggests that SUMO proteases require a self-conformational change prior to the cleavage reaction. Analysis of the interface of SENP1 and SUMO-1 has identified four unique amino acids in SENP1 that facilitate the interaction between SENP1 and SUMO-1. In addition to the structural study, functional characteristics of SENP1 have been analysed. In the present report, we show a novel function of SENP1 in hydrolysing the thioester bond in E1–SUMO and E2–SUMO complexes. This result suggests a new regulatory mechanism for the sumoylation pathway. Failure of SENP1^{C603S} in the maturation, de-sumoylation and cleavage of thioester linkage implies that the position of the sulphur atom of the active cysteine has to be re-orientated.

MATERIALS AND METHODS

Protein preparation

SENP1C and the precursor form of SUMO-1 (residues 1–101), were prepared as described previously [12]. For SENP1^{C603S}, the mutation C603S was introduced by PCR using the QuikChange[®] site-directed mutagenesis kit (Stratagene) and purified using the same procedure for wild-type SENP1 [12]. E1 conjugating ligase was expressed as a GST (glutathione S-transferase)–Uba2/Aos1 fusion protein. pGEX-Uba2/Aos1 clone was obtained from Professor Ron Hay (University of Dundee, Dundee, U.K.) and purified by glutathione–Sepharose affinity chromatography and Superdex[™] 200 gel-filtration (GE Healthcare). The E2 conjugating ligase, TDG (thymine DNA glycosylase; residues 112–360) was obtained from Dr Jane Wibley (Syngenta, Bracknell, U.K.) and Professor Laurence Pearl (Institute of Cancer Research, London, U.K.). Mature forms of SUMO-1 and SUMO-3 were expressed as N-terminal His₆-tag fusion proteins and purified by nickel affinity chromatography and Superdex[™] 75 gel-filtration (GE Healthcare). For the crystallization experiment, His₆-tagged truncated mature SUMO-1 (residues 20–97) was cloned and purified using the protocol as described above for the precursor form.

In vitro activity assays

TDG sumoylation reaction was performed by incubating TDG with E1, E2 and mature SUMO-1 or -3 at 37°C for 30 min with 10 mM ATP. The hydrolytic activity was investigated by incubating SENP1C or SENP1^{C603S} with sumoylated TDG, pre-matured SUMO or thioester linked E1–SUMO and E2–SUMO under normal conditions (20 mM Tris/HCl, pH 7.5 and 150 mM NaCl). All reactions were terminated by adding 5× protein loading dye [10% (w/v) SDS] and the reaction products were subjected to SDS/PAGE (12% gels) analysis and visualised with Coomassie Blue staining.

Purification and crystallization of SENP1^{C603S}–SUMO-1 complex

Purified SENP1^{C603S} was incubated with His₆-tagged SUMO-1 (residues 20–97) at 4°C for 1 h. SENP1^{C603S}–SUMO-1 complex was purified using a nickel–agarose column, followed by size exclusion on a Superdex[™] 75 column in buffer containing 10 mM Tris/HCl, pH 7.5 and 150 mM NaCl, to remove unbound SENP1C and SUMO-1. Initial crystallization screens were performed at 5 mg/ml with a microbatch method using Wizard screens I and II (Emerald Biosystems). Crystals grew after 3 days in a crystallization buffer {0.1 M Caps [3-(cyclohexylamino)propane-1-sulphonic acid], pH 10.5 and 40% (w/v) PEG [poly(ethylene glycol)] 400} at 16°C. This condition was optimized by mixing 1:1 (v/v) of protein complex [7.5 mg/ml in 10 mM Tris/HCl

Table 1 Diffraction data and refinement statistics of the SENP1^{C603S}–SUMO-1 complex

Values in parentheses are for the last resolution shell. $R_{\text{merge}} = (\sum h \sum j |I(h) - \langle I(h) \rangle| / \sum h \sum j \langle I(h) \rangle)$, where $\langle I(h) \rangle$ is the mean intensity of symmetry-equivalent reflections and $I(h)j$ is the measured intensity of the j th symmetry-related reflection. $R\text{-value} = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. Distribution of the residues in the most favoured/additionally allowed regions of the Ramachandran plot was evaluated by PROCHECK [26].

Parameter	Value
Diffraction data statistics	
Wavelength (Å)	1.54
Spacegroup	P4 ₃
Unit cell dimensions (Å)	97.31, 97.31, 102.58, $\alpha = 90^\circ$, $\beta = 90^\circ$, $\gamma = 90^\circ$
Resolution range (Å)	57.1–2.80 (2.95–2.80)
Number of observations	63395
Number of unique reflections	23472
Average redundancy	2.7 (2.7)
Completeness (%)	99.3 (99.9)
R_{merge}	0.11 (0.49)
Mean $I/\sigma I$	9.6 (2.0)
Refinement statistics	
Resolution	30–2.8
$R\text{-value}$	24.8
R_{free} (5% of data)	27.8
Number of atoms refined:	4662 (2 × 2331)
Protein (restrained)	
Water	54
Rmsd bond lengths (Å)	0.008
Rmsd bond angles (°)	1.27
Ramachandran plot most favoured/allowed region	87.8/12.2

(pH 7.5), 150 mM NaCl, 1 mM dithiothreitol and 0.5 mM EDTA] with precipitant containing 0.1 M Caps, pH 10.5, and a range of 30–40% PEG 400 using the hanging drop vapour diffusion procedure.

Data collection and structure determination

A 2.8 Å X-ray data set was collected at 100 K using a Rigaku MicroMax 007 X-ray generator (Centre for Protein Science and Crystallography, The Chinese University of Hong Kong, Hong Kong) and recorded on a R-Axis IV++ image plate (Rigaku). The crystallization buffer used was a cryoprotectant. Images were processed using MOSFLM [19], and scaled and reduced with SCALA (scale together multiple observations of reflections) from the CCP4 suite [20]. The protein complex crystallized in spacegroup P4₃, using the unit cell parameters: $a = 97.31$ Å, $b = 97.31$ Å, $c = 102.58$ Å, $\alpha = 90^\circ$, $\beta = 90^\circ$ and $\gamma = 90^\circ$. There were two complexes per asymmetric unit with a solvent content of approx. 57%. Statistics for the collected data are summarized in Table 1.

The structure was solved by molecular replacement using a molecule of SENP2–SUMO-1 (PDB ID code 1TGZ) [13] as a search model. The molecular replacement program in the CNS (crystallography and NMR system) was used with data in the resolution range 30–2.8 Å. After rigid-body refinement, the R - and R_{free} -factors of the initial solution were 39.9 and 39.5% respectively. It was followed by simulated annealing and positional refinements. Model building and refinement was performed using O [21] and CNS [22] with non-crystallographic symmetry restraints. Density modification and 2-fold map averaging were applied. The electron density maps from DM ($2F_o - F_c$) and ($F_o - F_c$) calculations were used for model building. The final model contained residues 439–644 of SENP1

Figure 1 De-conjugation and maturation assay for SENP1C and SENP1C^{C603S}

(a) SENP1C, but not SENP1C^{C603S}, is functional in de-sumoylation of SUMO-1 and SUMO-3 conjugated TDG. Sumoylation reactions were performed to conjugate SUMO-1 (lanes 1–3) or SUMO-3 (lanes 4–6) to TDG at 37 °C for 30 min by the SUMO conjugation system (0.5 μ g Uba2 and Aos1, 1 μ g Ubc9, 4 μ g SUMO, 10 mM ATP and 50 mM Tris/HCl, pH 7.5). After the conjugation reaction, 2 μ g SENP1C (lanes 2 and 5) or SENP1C^{C603S} (lanes 3 and 6) were incubated with the above sumoylation reaction mixtures for a further 10 min at 37 °C to perform the de-sumoylation reaction. TDG alone is shown in lane 7. (b) SUMO-1 maturation reaction catalysed by SENP1C and SENP1C^{C603S}. Purified SUMO-1–GST fusion protein (4 μ g) was incubated with 2 μ g of SENP1 (lane 1) or SENP1C^{C603S} (lane 2) or without SUMO-specific protease (lane 3) for 10 min at 37 °C. In both (a) and (b), the results were visualized by SDS/PAGE analysis.

and residues 20–97 of SUMO-1, in both complexes 54 water molecules with R -factor values of 24.8 % and R_{free} -factor values of 27.8 % were observed. Co-ordinates have been deposited in the Protein Data Bank (PDB ID 2G4D).

RESULTS AND DISCUSSION

Replacement of the active-site cysteine residue with serine in human SENP1 abolishes its hydrolytic activity for maturation and de-sumoylation

SENP1s play a role in both maturation and SUMO de-conjugation in the sumoylation pathway [12,13]. SENP1 exhibits hydrolytic activity in both maturation *in vitro* [12] and de-sumoylation *in vivo* [23]. Activity assays of the purified recombinant SUMO proteases showed that SENP1C, but not mutant SENP1C^{C603S}, was active in both desumoylation of SUMO-1 conjugated TDG (Figure 1a, lanes 1–3) and maturation of SUMO-1 (Figure 1b). The same phenomenon was observed in the SUMO-3 de-conjugation reactions (Figure 1a, lanes 4–6). No cleavage product was observed even in the presence of a high molar ratio (1:1, v/v) of SENP1C^{C603S} to substrates, indicating that substitution of cysteine at the active site of human SENP1C completely abolishes its catalytic activity. This result is consistent with the structural analysis of the Ulp1–SMT3 complex and mutagenesis of Ulp1 in yeast [18].

SENP1 is able to hydrolyse the thioester linkage in E1–SUMO and E2–SUMO complexes

To further characterize the biochemical activity of SENP1 in the sumoylation pathway, cleavage of the thioester linkage formed during SUMO activation was examined. When E1 was incubated with matured SUMO-1 in the presence of ATP, but not E2, a protein band corresponding to the ATP-dependent thioester complex Uba2–SUMO-1 was observed (Figure 2a, lane 2). This band disappeared after incubation with SENP1C, demonstrating that SENP1C displayed activity in thioester cleavage (Figure 2a, lane 3). Incubation of SUMO-3 instead of SUMO-1 also induced the formation of the thioester complex Uba2–SUMO-3, which could again be hydrolysed by the addition of SENP1C (Figure 2a, lanes 6 and 7). However, SENP1C^{C603S} did not exhibit any activity

toward the thioester cleavage for either SUMO-1 or -3 (Figure 2a, lanes 4 and 8 respectively).

To obtain more data on the observed thioesterase activity of SENP1C, the thioester linkage in E2–SUMO complex was studied. In the SUMO-conjugation pathway, transfer of activated SUMO from E1 to E2 results in a thioester complex between E2 and SUMO. In Figure 2(b), incubation of E1, E2 and SUMO in the presence of ATP induced the formation of a complex between E2 and SUMO (Figure 2b, lanes 2 and 6 correspond to E2–SUMO-1 and E2–SUMO-3 respectively). The observed protein band was clearly due to the accumulation of thioester-bond-linked E2–SUMO and not sumoylated E2. Addition of a large quantity of SUMO substrate, such as TDG or E1 (Uba2 of E1 is a SUMO substrate), caused this protein band to disappear (results not shown). As judged from Figure 2(b), lanes 3–7, SENP1C was also functional in cleaving the thioester linkage of the E2–SUMO complex.

Similar to the results of de-sumoylation and maturation assays shown in Figures 1 and 2, SENP1C^{C603S} was found to be inactive in all hydrolysis reactions (Figure 1a, lanes 3 and 6; Figure 1b, lane 3; Figures 2a and b, lanes 4 and 8). These results show that peptidase and thioesterase activities of SENP1C are catalysed by the same active-site cysteine residue with closely similar mechanisms. From lane 6 of Figure 2(b), it was also found that the SUMO-3 dimer was formed, which is approx. 2 kDa smaller than the E2–SUMO-3 complex. Our results suggest that SENP1C has an ability to hydrolyse SUMO-2 and SUMO-3 polymers (Figure 2b, lane 7).

We have demonstrated previously that SENP1 regulates the availability of different mature SUMOs by distinct maturation efficiencies [12]. The present findings suggest that SUMO proteases regulate the sumoylation pathway, not only during maturation and de-conjugation, but also in E1 activation and E2 conjugation processes. Our observations are in line with the fact that metabolic pathways are usually regulated in the early stages, so that they allow for an immediate and energy-saving response.

Structure of the SENP1C^{C603S}–SUMO-1 complex

To understand the molecular basis of catalytic mechanism of human SENP1 and its SUMOs specificities, we purified mutant SENP1C^{C603S} and truncated mature His₆-tagged SUMO-1

Figure 2 SENPI1C is functional in hydrolysis of the thioester linkage of E1-SUMO and E2-SUMO

(a) E1 (5 μ g of Uba2 and Aos1) was incubated with 5 μ g SUMO-1 (lanes 1–4) or SUMO-3 (lanes 5–8), in the presence of 10 mM ATP (lanes 2–4 and 6–8) or without (lanes 1 and 5) for 30 min at 37 °C. After incubation, 2 μ g SENPI1C (lanes 3 and 7) or SENPI1C^{C603S} (lanes 4 and 8) was added to the reaction mixture for a further 10 min at 37 °C. (b) E1 (0.5 μ g), E2 (5 μ g of Ubc9) and 5 μ g of SUMO-1 (lanes 1–4) or SUMO-3 (lanes 5–8) were incubated with (lanes 2–4 and 6–8) or without (lanes 1 and 5) 10 mM ATP for 30 min at 37 °C. After incubation, 2 μ g SENPI1C (lanes 3 and 7) or SENPI1C^{C603S} (lanes 4 and 8) were added to the reaction mixtures for a further 10 min at 37 °C. In both (a) and (b), the results were visualized by SDS/PAGE analysis.

(residues 20–97). The SENPI1C^{C603S}–SUMO-1 complex was produced by pull-down affinity chromatography and further purified by gel filtration. The purified SENPI1C^{C603S}–SUMO-1 complex was crystallized and the structure determined at 2.8 Å resolution (Table 1). Two complexes were present per asymmetric unit. The structure was solved by molecular replacement using SENP2–SUMO-1 [13] as a search model. Two copies of complexes present per asymmetric unit were refined by tight non-crystallographic symmetry restraints. The final model contained residues 439–644 of SENP1 and residues 20–97 of SUMO-1 with an R value of 24.8% and an R_{free} value of 27.8% (Figure 3). At the N-terminal of the SENPI1C^{C603S}, amino acids 427–438 have no discernible electron density and appear to be disordered. Both the structures of SENPI1C^{C603S} and mature SUMO-1 resemble those of SENP2 and SUMO-1, as described previously [13], comprising analogous secondary structure elements (Figure 4). From Figure 3, the N-terminal subdomain of SENPI1C consists of two β -strands (β 1 and β 2) and four α -helices (α 1, α 2, α 5 and α 6) while the C-terminal subdomain consists of a five stranded β -sheet (β 3– β 7) surrounded by two α -helices (α 3 and α 4). The protease catalytic triad Cys⁶⁰³ (mutated to serine in SENPI1C^{C603S}), His⁵³³ and Asp⁵⁵⁰ are located at the front of α 5, the front of β 5 and at the end of β 6 respectively. For SUMO-1, it is composed of the same

secondary-structure elements as shown in the SENP2–SUMO-1 complex. The extended C-terminal peptide terminated with the Gly-Gly motif binds to the catalytic pocket roofed by Trp⁴⁶⁵.

Using the DALI (distance matrix alignment) server [24] to search the Protein Data Bank to identify proteins structurally related to SENPI1C^{C603S}–SUMO-1 complex revealed the highest structural similarity with SENP2–SUMO-1, with a rmsd (root mean square deviation) of 0.9 Å over 205 aligned residues for SENP1 and a rmsd of 0.9 Å over 78 aligned residues for SUMO-1. Superimposition of the two complexes on the basis of SENP1 position shows that segments that align well correspond to residues 60–70 and 89–97 of SUMO, which are involved in protease interaction (Figure 7a). However, the SUMO core domain is relatively displaced from SENP2. Analysis of the intra-molecular interaction of the two structures shows that the distribution patterns of salt bridges are very similar. There are 12 salt bridges in the structure of SENPI1C^{C603S}, of which Asp⁵¹⁷ and Asp⁵⁷³ are variant, therefore the salt bridges between Asp⁵¹⁷ and Lys⁵⁷⁵, and Asp⁵⁷³ and Lys⁵⁷⁴ do not exist in the structure of SENP2. Of the 11 salt bridges in SENP2, the salt bridge between Glu³⁶⁷ and Lys⁵⁷⁸ connecting the N-terminal α 1 and α 7 helices could not be observed in SENP1, due to the disordered N-terminal of the protein.

Figure 3 Structure of SENP1^{C603S} catalytic domain in complex with SUMO-1

Stereo images of human SENP1^{C603S} (blue)–SUMO-1 (purple) complex. Residue C603S is shown as a ball-and-stick representation. Strands and helices are numbered for SENP1^{C603S} and are lettered for SUMO-1. All graphic images were prepared using the PyMOL Molecular Graphics System (DeLano Scientific).

Figure 4 Structure alignment of SENP1^{C603S} and SENP2

Conserved residues are shaded. The position of residue C603S is indicated by a star and the four unique residues 449, 455, 494, 514 involved in SENP1^{C603S}/SUMO-1 interaction are indicated by triangles and coloured in red and yellow respectively in Figure 5a. Residues conserved in the protease–SUMO interaction are highlighted by open circles and shown as green patches in Figure 5a. Electron density of residues 427–438 was not observed. Sequence numbering and secondary structure elements for SENP1 and SENP2 are shown above and below the alignment respectively. The Figure was produced using ESPript [27].

Interaction of SENP1 and SUMO-1

Analysis of the interaction surface of the SENP1^{C603S}–SUMO-1 complex revealed that in addition to the six interacting motifs found in the SENP2–SUMO-1 and Ulp1–Smt3 complexes, four unique residues in SENP1 that hydrogen-bonded with SUMO were identified, namely SUMO-Glu⁷⁹–Oδ...NH-SENPI-Arg⁴⁴⁹–NH...Oε-Asn⁷⁴–SUMO, SENP1-Lys⁴⁵⁵–Nζ...Oδ-Glu⁹³–SUMO, SENP1-Asn⁴⁹⁴–Nδ...O-Gly⁶⁸–SUMO and SENP1-Lys⁵¹⁴–Nζ...Oε-Glu⁶⁷–SUMO (Figure 5). These four residues in SENP1 are localized at the four corners of the interaction surface with the SUMO-1 core domain. Although Lys⁵¹⁴ is conserved in SENP2, the side chain of this amino acid points outward and forms a hydrogen bond with a water molecule instead of interacting with the SUMO molecule in the structure of SENP2. It is suggested that these extra unique hydrogen-bonded pairs found in the SENP1^{C603S}–SUMO-1 structure facilitate the binding of the SUMO-1 core domain.

The C-terminal extended peptide of SUMO-1 lies over the catalytic cleft through the hydrophobic tunnel formed by Trp⁴⁶⁵ and Trp⁵³⁴ as shown in the SENP2–SUMO-1 structure. By comparison with the SENP2 apo form and SENP2–SUMO-1 complex, it is interesting to note that although SENP1^{C603S} is inactive, the conformation of the Trp⁴⁶⁵, His⁵²⁹ and Trp⁵³⁴ residues resemble their equivalents in the SENP2–SUMO-1 complex (Figure 6). This implies that, consistent with SENP2, binding of SUMO to SENP1 triggers a local structural rearrangement of these residues. However, the imidazole ring of His⁵³³ is in a conformation that resembles the apo form of SENP2.

SENP1 requires a self-conformational change before the cleavage reaction

In the *in vitro* assays, mutant SENP1^{C603S} has no activity in the hydrolysis reactions. As the cleavage reaction of cysteine and serine proteases have very similar mechanisms (the OH-group

Figure 5 Interaction of SENP1C^{C603S} and SUMO-1

(a) Secondary structures of SENP1C^{C603S} and SUMO-1 are coloured in blue and purple respectively. Molecular surface of SENP1 is shown. Positions of the conserved residues involved in SUMO-1 interaction found in the SENP1C^{C603S}-SUMO-1 and SENP2-SUMO-1 [13] complexes are coloured in green. The four residues identified as unique in the SENP1C^{C603S}-SUMO-1 interaction are coloured in yellow and labelled. Residue C603S is coloured in red. Panels (b, c, d and e) show the four unique residues involved in the interaction between SENP1C^{C603S} and SUMO-1 using the ball-and-stick representation. The $(2F_o - F_c)$ electron density map contoured at 1-sigma level covering the interacting residues is shown

Figure 6 Structural superimposition of SENP1C^{C603S}-SUMO-1, SENP2-SUMO-1 and SENP2 apo enzyme

Stereo superimposition images of SENP1C^{C603S}-SUMO-1 (blue), SENP2-SUMO-1 (green) and SENP2 alone (magenta) on the basis of structural similarity analysed using the DALI server. Labelled residues are on the basis of the SENP1 sequence. Orientations of residues His⁴²⁹, Trp⁴⁶⁵ and Trp⁵³⁵ are similar to their equivalent residues in the SENP2-SUMO-1 complex, while His⁵³³ is more similar to its equivalent in the SENP2 apo enzyme. The 2.8 Å hydrogen bond, Ser⁶⁰³-OH...Nδ1-His⁵³³ is formed, which places Ser⁶⁰³ 4.1 Å away from the carbonyl-C of Gly⁹⁷ of SUMO-1.

of serine is even stronger than the SH-group of cysteine for hydrolysis activity), we did not know why the C603S mutation was devoid of activity.

It has been shown that SENP2 undergoes local structural rearrangements in response to SUMO-1 binding, which include reorientation of His⁴⁷⁸ towards Asp⁴⁹⁵ at the catalytic triad, rotation around the Cβ of Trp⁴¹⁰, Trp⁴⁷⁹ and His⁴⁷⁴ (in the SENP2 sequence) [13]. As described above and shown in Figure 6, only rearrangement of Trp⁴⁶⁵, His⁵²⁹ and Trp⁵³⁴ are observed in SENP1C^{C603S}-SUMO-1. The structure of SENP1C^{C603S}-SUMO-1 reveals that at the active site, a 2.83 Å hydrogen bond (Ser⁶⁰³-OH...Nδ1-His⁵³³) is formed, which is positioned far away (4.11 Å) from the carbonyl-C of Gly⁹⁷ of SUMO-1 (Figure 6). As the

hydrogen bond of OH...N is stronger than that of SH...N, this may explain why His⁵³³ cannot undergo a local rearrangement at the catalytic triad and therefore why the carbonyl-C of Gly⁹⁷ of SUMO-1 cannot be attacked.

In the structure of the SENP2-SUMO-1 complex, sodium borohydride was used to investigate the selective cleavage and reduction of the deacylation intermediate of the proteolytic reaction [13]. This indicates that the structural rearrangements of SENP2 facilitate either hydrolytic cleavage or release of SUMO. The structure of the SENP1C^{C603S}-SUMO-1 complex only follows part of such a rearrangement, which is induced by SUMO interaction, strongly suggesting that the active-site cysteine residue and the histidine of the catalytic triad in SUMO

Figure 7 Structural superimposition of SENP1C^{603S}–SUMO-1, SENP2–SUMO-1 and SENP1–SUMO-2

(a) Structural superimposition of SENP1C^{603S}–SUMO-1, SENP2–SUMO-1 and SENP1–SUMO-2 complexes. Stereo superimposition images of SENP1C^{603S}–SUMO-1 (blue), SENP2–SUMO-1 (green) and SENP1–SUMO-2 (orange) on the basis of structural alignment on protease molecules analysed using the DALI server. Labelled residues N⁶⁰, R⁷⁰, E⁸⁹ and G⁹⁷ highlight the regions of SUMO that are superimposed well among the three structures. The core domain of SUMO-2 shows the largest displacement from the protease. (b) Surface representation of the interaction of SUMO-1 with SENP1C^{603S}. Surface oxygen and nitrogen atoms of SENP1C are coloured in red and blue respectively. SUMO is shown in ball-and stick mode. Predicted acidic binding groove to SUMO-1, generated from backbone oxygen of Asn⁵⁹⁹ is indicated by a circle. (c) Shows the same as (b), but represents interaction of the SENP2–SUMO-1 complex. (d) Shows the same as (b), but represents interaction of the SENP1–SUMO-2 complex.

proteases need a self-conformational change before the cleavage reaction.

Taken together, the observations of the proteolytic mechanism of SUMO-specific proteases can be delineated. SENP1 and SENP2 appear catalytically competent in their apo forms. Upon SUMO interaction, local conformational rearrangement of Trp⁴⁶⁵, His⁵²⁹ and Trp⁵³⁴ occurs, while Cys⁶⁰³ separates from the catalytic triad to attack the carbonyl-C of Gly⁹⁷ of SUMO. After the cleavage event, SUMO is released from the complex and converted into its apo form ready for the next reaction cycle.

During the preparation of this manuscript, the structure of the human SENP1–SUMO-2 complex was published [25]. Structural comparison of the published SENP1–SUMO-2 with our own SENP1C^{603S}–SUMO-1 complex using the DALI server [24] revealed an rmsd of 0.6 Å over 205 aligned residues of SENP1 and an rmsd of 0.9 Å with 78 aligned residues of SUMO. Superimposition of the three available complex structures on the basis of protease structure (Figure 7a) clearly demonstrates the displacement of SUMO-2 from the protease as reported, which is found to be the largest among the three superimposed structures. Shen et al. [25] have suggested the relationship between maturation efficiency and electrostatic interactions of SUMO-specific proteases with the C-terminal peptides of SUMO precursors. However, our observation is different from their results by comparison of the electrostatic distributions of the SENP1C^{603S}–SUMO-1 complex with the SENP2–SUMO-1

complex (Figures 7b and 7c). The two structures do not show any strongly acidic or hydrophobic regions near the tunnel that would be complementary to the C-terminal peptides after the Gly-Gly cleavage site of SUMO. Instead, we observe a specific acidic binding groove generated from the backbone oxygen of Asn⁵⁹⁹ in the SENP1C^{603S}–SUMO-1 complex. Together with the biochemical data, it is most likely that the correct positioning of the two amino acid residues after the Gly-Gly motif would facilitate the hydrolysis reaction. From Figure 7(b), the highest substrate preference of SENP1 is to SUMO-1, which may be explained by the interaction of His⁹⁸ of SUMO-1 with the predicted binding groove. Nevertheless, we agree with Shen et al. [25] that Pro⁹⁴ of SUMO-3 would interfere with the binding to SENP1 and SENP2 due to the kink that the proline residue may introduce. In the present study, structural rearrangement of Trp⁴⁶⁵, His⁵²⁹ and Trp⁵³³ upon SUMO binding and orientation of the catalytic triad, Ser⁶⁰³, His⁵³³ and Asp⁵⁵⁰ observed in our SENP1C^{603S}–SUMO-1 structure provides insights into the molecular basis of SUMO binding and proteolysis. More importantly, our study demonstrates a novel function of SENP1 in thioester cleavage of E1–SUMO and E2–SUMO, which suggests an alternative regulatory mechanism that effectively modulates the SUMO activation and conjugation processes. In conclusion, it is clear that from the structures of SENP1, SENP2 and yeast Ulp1 in complex with SUMOs that SUMO proteases share a common catalytic mechanism. However, how SENP1 interacts with E1–SUMO and E2–SUMO is yet to be elucidated although studies

on the crystal structure of the thioester complex with SUMO-specific protease will resolve this. Finally, all the current structural analysis is limited to the catalytic domain, therefore investigation on how the non-conserved N-terminal domain participates in conjugated substrate specificity will be essential to unravel the de-sumoylation process.

This work was supported by the Competitive Earmarked Research Grant (CUHK 4409/04M), Research Grant Council, Hong Kong.

REFERENCES

- Verger, A., Perdomo, J. and Crossley, M. (2003) Modification with SUMO. *EMBO Rep.* **4**, 137–142.
- Muller, S., Ledl, A. and Schmidt, D. (2004) SUMO: a regulator of gene expression and genome integrity. *Oncogene* **23**, 1998–2008.
- Li, T. W., Evdokimov, E., Shen, R. F., Chao, C. C., Tekle, E., Wang, T., Stadtman, E. R., Yang, C. H. and Chock, P. B. (2004) Sumoylation of heterogeneous nuclear ribonucleoproteins, zinc finger proteins, and nuclear pore complex proteins: a proteomic analysis. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8551–8556.
- Denison, C., Rudner, A. D., Gerber, S. A., Bakalarski, C. E., Moazed, D. and Gygi, S. P. (2005) A proteomic strategy for gaining insights into protein sumoylation in yeast. *Mol. Cell. Proteomics* **4**, 246–254.
- Buschmann, T., Fuchs, S. Y., Lee, C. G., Pan, Z. Q. and Ronai, Z. (2000) SUMO-1 modification of Mdm2 prevents its self-ubiquitination and increases Mdm2 ability to ubiquitinate p53. *Cell* **101**, 753–762.
- Kadare, G., Toutant, M., Formstecher, E., Corvol, J. C., Carnaud, M., Bouterin, M. C. and Girault, J. A. (2003) PIAS1-mediated sumoylation of focal adhesion kinase activates its autophosphorylation. *J. Biol. Chem.* **278**, 47434–47440.
- Hietakangas, V., Ahlskog, J. K., Jakobsson, A. M., Hellesuo, M., Sahlberg, N. M., Holmberg, C. I., Mikhailov, A., Palvimo, J. J., Pirkkala, L. and Sistonen, L. (2003) Phosphorylation of serine 303 is a prerequisite for the stress-inducible SUMO modification of heat shock factor 1. *Mol. Cell. Biol.* **23**, 2953–2968.
- Swaminathan, S., Kiendl, F., Korner, R., Lupetti, R., Hengst, L. and Melchior, F. (2004) RanGAP1*SUMO1 is phosphorylated at the onset of mitosis and remains associated with RanBP2 upon NPC disassembly. *J. Cell Biol.* **164**, 965–971.
- Kamitani, T., Kito, K., Nguyen, H. P., Fukuda-Kamitani, T. and Yeh, E. T. (1998) Characterization of a second member of the sentrin family of ubiquitin-like proteins. *J. Biol. Chem.* **273**, 11349–11353.
- Bohren, K. M., Nadkarni, V., Song, J. H., Gabbay, K. H. and Owerbach, D. (2004) A M55V polymorphism in a novel *SUMO* gene (*SUMO-4*) differentially activates heat shock transcription factors and is associated with susceptibility to type I diabetes mellitus. *J. Biol. Chem.* **279**, 27233–27238.
- Owerbach, D., McKay, E. M., Yeh, E. T., Gabbay, K. H. and Bohren, K. M. (2005) A proline-90 residue unique to SUMO-4 prevents maturation and sumoylation. *Biochem. Biophys. Res. Commun.* **337**, 517–520.
- Xu, Z. and Au, S. W. (2005) Mapping residues of SUMO precursors essential in differential maturation by SUMO-specific protease, SENP1. *Biochem. J.* **386**, 325–330.
- Reverter, D. and Lima, C. D. (2004) A basis for SUMO protease specificity provided by analysis of human Senp2 and a Senp2–SUMO complex. *Structure* **12**, 1519–1531.
- Boggio, R., Colombo, R., Hay, R. T., Draetta, G. F. and Chiocca, S. (2004) A mechanism for inhibiting the SUMO pathway. *Mol. Cell* **16**, 549–561.
- Murtas, G., Reeves, P. H., Fu, Y. F., Bancroft, I., Dean, C. and Coupland, G. (2003) A nuclear protease required for flowering-time regulation in *Arabidopsis* reduces the abundance of small ubiquitin-related modifier conjugates. *Plant Cell* **15**, 2308–2319.
- Kim, K. I., Baek, S. H. and Chung, C. H. (2002) Versatile protein tag, SUMO: its enzymology and biological function. *J. Cell. Physiol.* **191**, 257–268.
- Kim, Y. H., Sung, K. S., Lee, S. J., Kim, Y. O., Choi, C. Y. and Kim, Y. (2005) Desumoylation of homeodomain-interacting protein kinase 2 (HIPK2) through the cytoplasmic-nuclear shuttling of the SUMO-specific protease SENP1. *FEBS Lett.* **579**, 6272–6278.
- Mossessova, E. and Lima, C. D. (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol. Cell* **5**, 865–876.
- Leslie, A. (1995) *MOSFLM Users Guide*, MRC Laboratory of Molecular Biology, Cambridge, U.K.
- Collaborative Computational Project, Number 4 (1994) The *CCP4* suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763.
- Jones, T. A., Zou, J. Y., Cowan, S. W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **43**, 110–119.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Grow, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S. et al. (1998) Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921.
- Bailey, D. and O'Hare, P. (2004) Characterization of the localization and proteolytic activity of the SUMO-specific protease, SENP1. *J. Biol. Chem.* **279**, 692–703.
- Holm, L. and Sander, C. (1996) Alignment of three-dimensional protein structures: network server for database searching. *Methods Enzymol.* **266**, 653–662.
- Shen, L., Dong, C., Liu, H., Naismith, J. H. and Hay, R. T. (2006) The structure of SENP1 SUMO-2 co-complex suggests a structural basis for discrimination between SUMO paralogs during processing. *Biochem. J.* **397**, 279–288.
- Laskowski, R. A., MacArthur, M. W. and Thornton, J. M. (1998) Validation of protein models derived from experiment. *Curr. Opin. Struct. Biol.* **8**, 631–639.
- Gouet, P., Courcelle, P., Stuart, D. I. and Metz, F. (1999) ESPript: multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308.

Received 10 April 2006/17 May 2006; accepted 22 May 2006

Published as BJ Immediate Publication 22 May 2006, doi:10.1042/BJ20060526