Vine data

461

## Activity profiling of human deSUMOylating enzymes (SENPs) with synthetic substrates suggests an unexpected specificity of two newly characterized members of the family

Marcin DRAG\*+1, Jowita MIKOLAJCZYK\*, I. M. KRISHNAKUMAR‡, Ziwei HUANG‡ and Guy S. SALVESEN\*1

\*Apoptosis and Cell Death Research Program, Burnham Institute for Medical Research, La Jolla, CA 92037, U.S.A., †Division of Medicinal Chemistry and Microbiology, Faculty of Chemistry, Wroclaw University of Technology, Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland, and ‡Signal Transduction Program, Burnham Institute for Medical Research, La Jolla, CA 92037, U.S.A.

SENPs [Sentrin/SUMO (small ubiquitin-related modifier)specific proteases] include proteases that activate the precursors of SUMOs, or deconjugate SUMOs attached to target proteins. SENPs are usually assayed on protein substrates, and for the first time we demonstrate that synthetic substrates can be convenient tools in determining activity and specificity of these proteases. We synthesized a group of short synthetic peptide fluorogenic molecules based on the cleavage site within SUMOs. We demonstrate the activity of human SENP1, 2, 5, 6, 7 and 8 on these substrates. A parallel positional scanning approach using a fluorogenic tetrapeptide library established preferences of SENPs in the P3 and P4 positions that allowed us to design

# optimal peptidyl reporter substrates. We show that the specificity of SENP1, 2, 5 and 8 on the optimal peptidyl substrates matches their natural protein substrates, and that the presence of the SUMO domain enhances catalysis by 2–3 orders of magnitude. We also show that SENP6 and 7 have an unexpected specificity that distinguishes them from other members of the family, implying that, in contrast to previous predictions, their natural substrate(s) may not be SUMO conjugates.

Key words: combinatorial chemistry, fluorogenic substrate library, protease, SENP, SUMO.

#### INTRODUCTION

The human peptidase clan CE contains 7 family members, also known as SENPs [Sentrin/SUMO (small ubiquitin-related modifier)-specific proteases] [1,2]. These enzymes play a similar role to deubiquitinating enzymes by participating in the control of small ubiquitin-like modifiers. In humans, small ubiquitinlike modifiers consist of at least 12 members, among which the best characterized is the SUMO family [3-5]. Recent studies suggest that SUMO-1, -2 and -3 are putative natural substrates of most SENPs. To date, SUMO-4 is considered a structural homologue of SUMO-1, -2 and -3, but it is not clear that it is a SENP substrate. SENP8 has less similarity with other family members and demonstrates activity only against Nedd8-modified proteins [6]. There are two major proteolytic events in which SENPs participate. First is the maturation of pro-SUMOs or pro-Nedd8 by removal of C-terminal tails to uncover a Gly-Gly motif of the C-terminus — the endopeptidase function, required for subsequent ligation to their target proteins. Second is the removal of SUMOs or Nedd8 from the target proteins - the isopeptidase, or C-terminal hydrolase, function [3,7]. On the basis of protein sequence similarity, human SENPs can be divided into four groups, with SENP1 and 2 forming one group, SENP3 and 5 another group, SENP6 and 7 another, and SENP8 a group of its own [8]. It is generally assumed that, with the exception of SENP8, all other SENPs cleave SUMOs, although this has only been investigated in detail for SENP1, 2, 3 and 5 [9–11].

SENPs are usually assayed by their activity on protein substrates, which restricts the type of analyses that can be performed [9]. Structural data reveal two regions that dictate SENP specificity: the residues lining the active site cleft that recognize the four exposed C-terminal residues, and the extended interface that recognizes the SUMO fold (an exosite interaction). Small synthetic substrates have been used previously to investigate DUBs (deubiquitinating enzymes), a distantly related and much more extensively investigated family of proteases, but catalysis was substantially depressed compared with the natural substrate [12]. In the case of SENPs, the only report on small molecules comes from the observation that a pentapeptide vinyl sulfone probe based on the C-terminal sequence of SUMO-1 labels some proteases in cell lysates with a mass consistent with SENPs [13]. To understand the relative importance of the active site cleft compared with the exosite interaction in SENP activity and specificity, we have determined the minimal peptide substrate that can be cleaved by SENPs. We present the first synthetic approach and use of a PS-SCL (positional scanning substrate combinatorial library), in the form of tetrapeptides with an ACC (7-amino-4-carbamovlmethylcoumarin) fluorophore as leaving group, to define the specificity of the SENP active site cleft. Our results reveal stringent specificity for all human SENPs, with the exception of SENP3, confirming some predicted specificities, and also some totally unexpected specificities for SENP6 and 7.

#### **EXPERIMENTAL**

#### Materials

Chemicals and solvents were obtained from commercial suppliers and used without further purification, unless otherwise stated. Safety Catch and Rink amide resins were purchased from

Abbreviations used: Ac, acetyl; ACC, 7-amino-4-carbamoylmethylcoumarin; AFC, 7-amino-4-trifluoromethylcoumarin; Boc, *tert*-butoxycarbonyl; Cbz, benzyloxycarbonyl; DTT, dithiothreitol; DUB, deubiquitinating enzyme; ESI, electrospray ionization; IPTG, isopropyl-β-D-galactopyranoside; PS-SCL, positional scanning substrate combinatorial library; SUMO, small ubiquitin-related modifier; SENP, Sentrin/SUMO-specific protease; TFA, trifluoroacetate. <sup>1</sup> Correspondence may be addressed to either of these authors (email mdrag@burnham.org or gsalvesen@burnham.org)

Novabiochem. Anhydrous DMF (*N*,*N*-dimethyl formamide) was from Sigma–Aldrich. Human ubiquitin was purchased from Boston Biochem.

#### **Plasmid constructs**

Catalytic domains of SENP2 (residues 363-589), SENP7 (residues 640–984), and SENP8 (residues 1–212) were amplified from a human fetal brain cDNA library, and the catalytic domains of SENP1 (residues 419-643), SENP5 (residues 536-755) and SENP6 (residues 628-1112) were amplified from a human keratinocyte cDNA library as previously described in [14]. The PCR products were cloned into the bacterial expression vector pET28a (Novagen), engineered to contain an N-terminal Histag. C-terminally truncated SUMO and Nedd8 constructs were generated from full-length constructs and inserted into the pET28a vector in-frame with an N-terminal His-tag [14]. Using the wildtype cDNA as a PCR template, the following mutations of residues at the P4 and P3 positions were introduced: pro-SUMO-1 (single substitution at P3 of Thr<sup>95</sup> to Tyr, Ile or Gly) and proSUMO-2 (double mutant at P4 and P3 of Gln<sup>90</sup>-Thr<sup>91</sup> to Leu-Arg). Mutants were amplified with reverse primers designed to carry specific mutations. All constructs were cloned into the pET28a plasmid to generate proteins with a C-terminal His-tag.

#### Cleavage of proSUMOs by SENPs

To assay the processing of SUMO precursors, recombinant proSUMOs (5  $\mu$ M) were incubated with the indicated concentration range of recombinant SENPs for 1 h at 37 °C. Cleavage products were analysed using a 12% ammediol/HCl SDS/PAGE gel system and stained with GelCode Blue.

#### Protein expression in Escherichia coli

Recombinant SENP enzymes, SUMO proteins, and Nedd8 were produced in *E. coli* codon plus (Novagen). Production of SUMO proteins and Nedd8 was induced with 0.4 mM IPTG (isopropyl- $\beta$ -D-galactopyranoside) at 37 °C for 3 h. Expression of SENPs was induced with 0.2 mM IPTG at 30 °C for 3 h. His-tagged proteins were purified using Ni-NTA agarose and eluted with a 20–200 mM gradient of imidazole in 50mM Hepes (pH 7.4) plus 100 mM NaCl. Protein purity was confirmed by SDS/PAGE and concentrations of the purified proteins were determined from the absorbance at 280 nm based on the molar absorption coefficients determined from the Edelhoch relationship [15].

### Synthesis of the individual AFC (7-amino-4-trifluoromethylcoumarin) substrates

Both solution-phase synthesis and solid-phase synthesis were carried out according to well-established methods. In the solution-phase synthesis, the coupling of Boc-Gly-COOH with AFC and peptide chain elongation using a Boc (*tert*-butoxycarbonyl) strategy was based on a previous procedure [16]. Solid-phase synthesis was performed using Safety Catch resin exactly as described by Backes and Ellman [17]. Synthesis was performed using the semiautomatic *Flex*Chem Peptide Synthesis System (Model 202).

Post-synthesis, column chromatography was performed with grade 60 silica gel (Fisher, 70–230 mesh). Analytical HPLC analyses were conducted on a Beckman-Coulter System Gold 125 solvent delivery module equipped with a Beckman-Coulter System Gold 166 Detector system using a Varian Microsorb-MV C18 ( $250 \times 4.8$  mm) column. Preparative HPLC analysis were conducted on a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery module equipped with a Beckman-Coulter System Gold 126P solvent delivery Method System Gold 126P solvent delivery module equipped with Beckman-Coulter System Gold 126P solvent delivery module equipped with Beckman-Coulter System Gold 126P s

System Gold 168 Detector system using a Kromasil 100-10 C18 (20 mm ID) column (Richard Scientific). Solvent composition: system A [water/0.1% TFA (trifluoroacetate)] and system B [acetonitrile/water 80%:20% (v/v) with 0.1% of TFA]. Selected substrates were validated by MS recorded in ESI (electrospray ionization) mode with the aid of the Burnham Proteomics facility.

#### <sup>1</sup>H-NMR analysis of the individual AFC substrates

<sup>1</sup>H-NMR spectra were obtained with the aid of the Burnham Structural Biology facility using a Varian 300 spectrometer in [<sup>2</sup>H]chloroform or [<sup>2</sup>H<sub>6</sub>]DMSO (Aldrich). <sup>1</sup>H-NMR (300 MHz) spectra are reported as follows: chemical shifts in ppm downfield from TMS (tetramethylsilane), the internal standard; resonance signal description (b, broad; d, doublet; m, multiplet; s, singlet; t, triplet), integration, and coupling constant (Hz). ×TFA, one molecule of TFA per molecule of compound. Aminoprotecting free radicals not previously defined: Ac, acetyl; Cbz, benzyloxycarbonyl.

Boc-Gly-AFC: <sup>1</sup>H-NMR ([<sup>2</sup>H]chloroform): 1.44 (s, 9 H), 3.90 (d, 2 H, J 5.7 Hz), 5.21 (bs, 1 H), 6.63 (s, 1 H), 7.28 (d, 1 H, J 8.0 Hz), 7.47 (d, 1 H, J 8.0 Hz), 7.50 (s, 1 H), 8.82 (bs, 1 H); Boc-Gly-Gly-AFC: <sup>1</sup>H-NMR ([<sup>2</sup>H<sub>6</sub>]DMSO): 1.37 (s, 9 H), 3.59 (d, 2 H, J 5.1 Hz), 3.94 (d, 2 H, J 5.4 Hz), 6.89 (s, 1 H), 7.06 (s, 1 H), 7.53 (d, 1 H, J 8.1 Hz), 7.67 (d, 1 H, J 7.8 Hz), 7.89 (s, 1 H), 8.21 (s, 1 H), 10.55 (s, 1 H); Boc-Thr-Gly-Gly-AFC: <sup>1</sup>H-NMR ([<sup>2</sup>H<sub>6</sub>]DMSO): 1.05 (d, 3 H, J 5.9 Hz), 1.39 (s, 9 H), 3.81 (d, 2 H, J 5.4 Hz), 3.95 (m, 3 H), 4.86 (s, 1 H), 6.48 (d, 1 H, J 6.6 Hz), 6.92 (s, 1 H), 7.56 (d, 1 H, J 8.7 Hz), 7.69 (d, 1 H, J 8.3 Hz), 7.93 (s, 1 H), 8.29 (s, 2 H), 10.44 (s, 1 H); Boc-Gln-Thr-Gly-Gly-AFC: <sup>1</sup>H-NMR ([<sup>2</sup>H<sub>6</sub>]DMSO): 1.05 (d, 3 H, J 6.6 Hz), 1.39 (s, 9 H), 1.72 (m, 1 H), 1.88 (m, 1 H), 2.11 (m, 2 H), 3.83 (m, 2 H), 3.97 (m, 3 H), 4.24 (m, 1 H), 5.04 (d, 1 H, J 4.2 Hz), 6.77 (s, 1 H), 6.93 (s, 1 H), 7.15 (d, 1 H, J 7.2 Hz), 7.27 (s, 1 H), 7.55-7.73 (m, 3 H), 7.92 (s, 1 H), 8.23 (m, 1 H), 8.32 (m, 1 H), 10.51 (s, 1 H); Cbz-Gln-Thr-Gly-Gly-AFC: <sup>1</sup>H-NMR ([<sup>2</sup>H<sub>6</sub>]DMSO): 1.05 (d, 3 H, J 5.7 Hz), 1.74 (m, 1 H), 1.91 (m, 1 H), 2.14 (m, 2 H), 3.83 (d, 2 H, J 4.2 Hz), 3.98 (d, 2 H, J 5.8 Hz), 4.06 (m, 2 H), 4.25 (m, 1 H), 5.04 (s, 3 H), 6.78 (s, 1 H), 6.93 (s, 1 H), 7.27-7.36 (m, 5 H), 7.55–7.78 (m, 4 H), 7.92 (d, 1 H, J 0.9 Hz), 8.23 (m, 1 H), 8.34 (m, 1 H), 10.52 (s, 1 H); Ac-Gln-Thr-Gly-Gly-AFC: <sup>1</sup>H-NMR ([<sup>2</sup>H<sub>6</sub>]DMSO): 1.05 (d, 3 H, J 6.0 Hz), 1.72 (m, 1 H), 1.87 (m, 4 H), 2.09 (m, 2 H), 3.82 (d, 2 H, J 5.7 Hz), 3.98 (d, 2 H, J 5.4 Hz), 4.04 (m, 1 H), 4.20–4.31 (m, 2 H), 5.04 (bs, 1 H), 6.77 (s, 1 H), 6.93 (s, 1 H), 7.28 (s, 1 H), 7.57 (d, 2 H, J 8.7 Hz), 7.26 (m, 2 H), 7.93 (s, 1 H), 8.14 (d, 1 H, J 7.8 Hz), 8.22 (s, 1 H), 8.32 (s, 1 H), 10.50 (s, 1 H); Ac-Gln-Tyr-Gly-Gly-AFC: <sup>1</sup>H-NMR ([<sup>2</sup>H<sub>6</sub>]DMSO): 1.64 (m, 1 H), 1.83 (m, 4 H), 2.04 (m, 2 H), 3.78 (m, 2 H), 3.98 (d, 2 H, J 6.0 Hz), 4.19 (m, 1 H), 4.43 (m, 1 H), 6.62 (d, 2 H, J 7.8 Hz, AB), 6.76 (s, 1 H), 6.92 (s, 1 H), 7.01 (d, 2 H, J 8.4 Hz, AB), 7.24 (s, 1 H), 7.57 (d, 1 H, J 8.1 Hz), 7.93-8.02 (m, 3 H), 8.24 (m, 1 H), 8.37 (m, 1 H), 9.17 (s, 1 H), 10.49 (s, 1 H); Ac-Leu-Arg-Gly-Gly-AFC  $\times$  TFA: <sup>1</sup>H-NMR ([<sup>2</sup>H<sub>6</sub>]DMSO): 0.83 (m, 6 H), 1.40–1.73 (m, 7 H), 1.85 (s, 3 H), 3.11 (m, 2 H), 3.81 (m, 2 H), 3.98 (d, 2 H, J 6.0 Hz), 4.25–4.28 (m, 2 H), 6.94 (s, 1 H), 6.82-7.43 (2 × bs, 2 H) 7.49 (m, 1 H), 7.58 (d, 1 H, J 9.5 Hz), 7.72 (d, 1 H, J 7.5 Hz), 7.95 (d, 1 H, J 1.8 Hz), 8.04 (d, 1 H, J 7.8 Hz), 8.04 (d, 1 H, J 7.8 Hz), 8.18 (d, 1 H, J 6.9 Hz), 8.33–8.36 (m, 2 H), 10.44 (s, 1 H); Ac-Ala-Leu-Arg-Gly-Gly-AFC × TFA: <sup>1</sup>H-NMR ([<sup>2</sup>H<sub>6</sub>]DMSO): 0.83 (m, 6 H), 1.18 (d, 3 H, J 3.9 Hz), 1.45-1.78 (m, 7 H), 1.84 (s, 3 H), 3.10 (m, 2 H), 3.82 (m, 2 H), 3.98 (d, 2 H, J 5.1 Hz), 4.23–4.30 (m, 3 H), 6.94 (s, 1 H), 6.73–7.45 (2 × bs, 2 H) 7.46 (m, 1 H), 7.59 (d, 1 H, J 10.2 Hz), 7.72 (d, 1 H, J 7.2 Hz), 7.94 (d, 1 H, J 1.8 Hz), 8.02 (d, 1 H, J 7.5 Hz), 8.07 (d, 1 H, J 7.2 Hz), 8.35 (d, 2 H, J 5.7 Hz), 10.44 (s, 1 H); Ac-Gln-Gln-Thr-Gly-Gly-AFC: <sup>1</sup>H-NMR ([<sup>2</sup>H<sub>6</sub>]DMSO): 1.05 (d, 3 H, J 6.0 Hz), 1.71–1.88 (m, 7 H), 2.09 (m, 4 H), 3.81 (d, 2 H, J 5.1 Hz), 3.97 (d, 2 H, J 5.7 Hz), 4.04 (m, 1 H), 4.22–4.31 (m, 3 H), 6.78 (s, 2 H), 6.92 (s, 1 H), 7.26 (s, 2 H), 7.57 (d, 1 H, J 9.0 Hz), 7.70–7.77 (m, 2 H), 7.93 (s, 1 H), 8.09 (d, 1 H, J 7.5 Hz), 8.17–8.34 (m, 3 H), 10.50 (s, 1 H); Ac-Glu-Gln-Thr-Gly-Gly-AFC: <sup>1</sup>H-NMR ( $[^{2}H_{6}]DMSO$ ): 1.05 (d, 3 H, J 5.4 Hz), 1.73–1.86 (m, 7 H), 2.10 (m, 2 H), 2.25 (m, 2 H), 3.82 (d, 2 H, J 5.4 Hz), 3.98 (d, 2 H, J 5.7 Hz), 4.04 (m, 1 H), 4.24–4.33 (m, 3 H), 5.08 (bs, 1 H), 6.78 (s, 1 H), 6.93 (s, 1 H), 7.27 (s, 1 H), 7.57 (d, 1 H, J 9.0 Hz), 7.72 (d, 1 H, J 7.8 Hz), 7.84 (d, 1 H, J 5.6 Hz), 7.93 (s, 1 H), 8.09 (d, 1 H, J 6.2 Hz), 8.17–8.34 (m, 3 H), 10.49 (s, 1 H).

#### Assay of individual fluorogenic substrates

All individual substrates were screened against SENPs at 37 °C in low salt Tris buffer [50 mM Tris (pH 8.0), 20mM NaCl and 5mM DTT (dithiothreitol)]. Enzymes were preincubated for 10 min at 37 °C before adding substrate to the wells of an fMax fluorimeter (Molecular Devices) 96-well plate reader operating in the kinetic mode. Enzyme assay conditions were as follows (100  $\mu$ l reaction): 20  $\mu$ M final substrate concentration and enzymes at 0.25–4  $\mu$ M. Release of AFC fluorophore was monitored continuously with excitation at 405 nm and emission at 510 nm. Each experiment was repeated at least three times and the results presented as an average. Final substrate concentrations for  $k_{\text{cat}}/K_{\text{m}}$  determination ranged from 50–1  $\mu$ M. Concentration of DMSO in the assay was less than 1 % (v/v). To determine the catalytic efficiency of enzyme the initial velocities  $(v_i)$  were measured as a function of  $[S_0]$  (substrate concentration at zero time). When  $[S_0] \ll K_m$ , the plot of V versus  $[S_0]$  yields a straight line with slope representing  $V_{\text{max}}/K_{\text{m}}$ . The  $k_{\text{cat}}/K_{\text{m}}$  ratios were calculated using the following expression:

 $k_{\rm cat}/K_{\rm m} = {\rm slope}/{\rm E}$ 

where E is final enzyme concentration. We were unable to saturate the enzyme with substrate and therefore could not determine individual  $k_{cat}$  and  $K_m$  values.

#### Synthesis of the diverse tetrapeptide-ACC PS-SCL

Synthesis of the PS-SCL was carried out with the help of the Burnham Institute Peptide Synthesis Facility and is based on related procedures [18,19]. After completing synthesis and analysis steps, each sub-library was dissolved in biochemical-grade dried DMSO at a concentration of 10 mM and stored at -20 °C until use.

#### Assay of the PS-SCL

All the SENPs were assayed at 37 °C in an appropriate buffer system of either low salt Tris buffer, or sodium citrate buffer [25 mM Tris (pH 8.0), 0.8 M sodium citrate and 5 mM DTT]. Experiments with truncated SUMOs were carried out in the low salt Tris buffer at a 1:5 molar ratio of SENP to SUMO. All the enzymes were preincubated for 10 min at 37 °C before being added to the wells containing substrate. Standard enzyme assay conditions were as follows (100  $\mu$ l reaction). 250 or 500  $\mu$ M (assuming around 13 or 26  $\mu$ M per single substrate) total final substrate mixture concentration and enzyme concentration was 0.5–6.0  $\mu$ M. Release of fluorophore was monitored continuously with excitation at 355 nm and emission at 460 nm with an assay time of 30 min. Analysis of the results was based on total RFU (relative fluorescence unit) for every sub-library, setting the highest value to 100 % and adjusting the other results accordingly.

#### RESULTS

#### Synthesis and characterization of single substrates

All individual substrates were synthesized using either solutionor solid-phase chemistry. The preliminary step in both cases involved synthesis of Boc-Gly-AFC by attaching the fluorophore (AFC) to the protected Gly in pyridine and  $POCl_3$  as previously described in [16]. After purifying by column chromatography on silica gel, Boc-Gly-AFC was deprotected in 4 M HCl in dioxane. The yield after two steps was usually around 60%. In a solution-phase approach, elongation of the peptide chain was performed using a Boc strategy and HBTU/DIEA (O-benzotriazole-N,N,N',N'-tetramethyl-uroniumhexafluoro-phosphate/di-isopropylethylamine). After synthesis, the peptides were deprotected using TFA in the presence of scavengers, and purified using preparative HPLC. The final overall yield after all steps depended on the length of the peptide chain, and was usually 40-55%. This approach was applied to the synthesis of all substrates with a Boc- or Cbz-protecting group at the N-terminus. In the case of the solid-phase approach, the synthesis was carried out using a Safety Catch resin as described earlier in [17]. This protocol was applied for the synthesis of all N-terminal acetyl-protected peptides. The overall yield after synthesis and preparative HPLC purification was between 25 and 70% depending on the level of substitution of the resin. All the substrates were characterized by <sup>1</sup>H-NMR spectroscopy, confirming the structure, and were usually more than 95% pure as determined by analytical HPLC.

#### Search for the shortest active fluorogenic substrate

To define the shortest amino acid sequence upon which SENPs are reasonably active, and with which we can explore the specificity requirements within the SENP active site cleft, we synthesized substrates based on the C-terminal sequences of the presumed natural substrates: SUMO-1, -2 and -3 for SENP1, 2 and 5 respectively, and Nedd8 for SENP8. We used recombinant enzymes produced as catalytic domains in E. coli. SENP1, 2 and 5 demonstrated practically no activity on substrates containing one, two or three residues, and SENP8 showed minimal activity on a Cbz-RGG-AFC substrate (Figure 1). However, a dramatic increase in activity is observed after addition of a fourth residue, specifically the QTGG recognition element for SENP1, 2 and 5, and LRGG for SENP8 - sequences that correspond with the terminal tetrapeptides of SUMO or Nedd8 respectively. SENP1, 2 and 5 were completely inactive on the tetrapeptide PTGG substrate based on the SUMO-4 sequence (results not shown). Significantly, in the case of SENP1, 2 and 5, the protecting group at the N-terminus plays a substantial role, with an acetyl group allowing substantially higher substrate activity than Boc (Figure 1) or Cbz (results not shown). This is probably due to the bulky size of the latter protecting groups, which may disrupt the interaction of the substrate with recognition elements of SENP1, 2 and 5, or possibly a difference in hydrogen bonding resulting from the change in the electronic environment from an N-terminal carbamate to an N-terminal amide. However, this effect is not observed in the case of SENP8, which reveals almost the same substrate activity in the case of acetyl or Cbz protecting groups, as revealed by comparison of the activity of the commercially available DUB substrate Cbz-LRGG-AMC and our synthesized Ac-LRGG-AFC compound (Figure 1). Interestingly, SENP6 and 7 showed almost no activity on Ac-QTGG-AFC [14].

#### Synthesis of a tetrapeptide-ACC positional scanning library

To address the substrate specificity of the SENPs, we designed and synthesized a combinatorial library of tetrapeptide substrates



Figure 1 SENP1, 2, 5 and 8 are active on tetrapeptide substrates

Individual recombinant human SENP catalytic domains were expressed and purified from *E. coli.* Activity was screened using peptidyl AFC fluorogenic substrates based on the sequences encompassing the cleavage sites in SUMOs. The grey bar represents a comparison of the commercially available AMC-based fluorophore. The *y*-axis represents the activity as an average of at least three individual assays. The enzyme concentration was 4  $\mu$ M (based on total protein), substrate concentration 100  $\mu$ M, assay time 20 min. Rates are expressed as nM AFC product/min per  $\mu$ M enzyme.

with ACC as the fluorophore. In the design of the library, we fixed the P1 and P2 residues as Gly-Gly since these are strongly conserved among DUB and SENP natural substrates, and because examination of SENP crystal structures indicates that amino acids containing a side-chain are unlikely to be tolerated in the P1 and P2 positions [20–22]. Design of the library is presented in Figure 2. Our approach generates two libraries (P3 and P4 position) each containing 19 sub-libraries with an equimolar mixture of 19 sequences per sub-library, giving a total of 361 substrates. The advantage of the library in this form is that we can see strong fluorescence signals due to the presence of only 19 individual tetrapeptidic substrates in each sub-library, which is an advantage when specificity may be strict and catalytic rates are relatively low. Cysteine and methionine residues were omitted because they are usually considered problematic to handle due to oxidation. We substituted the methionine residue with norleucine, which is very similar in structure and differs only by one atom, namely a sulfur-to-carbon substitution at the  $\gamma$ -position. The N-terminus of the peptide chain is acetylated, as this was found to be the most suitable protecting group in earlier experiments. The synthesis of the library was performed as described previously for substrates of cathepsins and caspases [18,19,23]. In our case, the first two Gly residues were attached to the ACC fluorophore in one reaction vessel in order to obtain exactly the same level of substitution. The level of substitution of the first Gly residue to the ACC fluorophore, which is usually considered as the most problematic, was essentially 100%, as judged by analytical HPLC. After these



Figure 2 Schematic representation of the PS-SCL

Each library contains a mixture of 19 amino acids, omitting cysteine and methionine and including norleucine. The substrates were dissolved at 10 mM in DMSO and used at a final concentration of 250 or 500  $\mu$ M (representing 13 or 26  $\mu$ M per individual substrate sequence).

steps, the resin was dried, the level of substitution was calculated and parallel solid-phase synthesis was carried out. The presence of the expected substrates was validated by ESI MS of 5 randomly selected samples from each sub-library.

#### Results from screening of the positional scanning library

Our separate studies demonstrated that the activities of SENP1, 2, 5, 6 and 7 are sensitive to salt conditions, and can be enhanced by the addition of C-terminally truncated SUMOs [14]. Therefore we scanned SENP1, 2, 5, 6 and 7 under diverse conditions, with the aim of determining whether substrate specificity was altered under these conditions. Addition of truncated SUMOs or use of high concentrations of sodium citrate enhanced catalysis but did not substantially influence substrate specificity (Supplemental Figure 1, http://www.BiochemJ.org/bj/409/bj4090461add.htm), so we present in Figure 3 specificity comparisons based on optimal assay conditions.

We immediately recognized three substrate preference groupings, which we term groups I, II and III. Group I, which contains SENP1, 2 and 5, reveals high specificity for the glutamine residue in P4 and significant, but much lower, stringency in P3. At the P3 position, both SENP1 and 2 reveal much greater tolerance for almost all amino acids with the notable exceptions of glycine and proline. For SENP5, we performed assays only in the presence of sodium citrate conditions, because attempts to determine preference in low salt conditions or in the presence of truncated SUMOs failed due to low activity of the enzyme. SENP5 was a little more restricted at P3 since it preferred tyrosine, but also tolerated most other residues to some extent. Truncated SUMOs or sodium citrate did not substantially alter the wide tolerance at P3, with the exception that tyrosine now became the amino acid preferred over threonine (Supplemental Figure 1). For the most part, the substrate preference matches that of SUMO-1, -2 and -3 (QTGG), and the high preference in the P4 position and relatively low specificity in the P3 position can be explained by comparison with the published crystal structures of SENP1 and 2 with SUMOs (Figure 4) [20-22]. The side-chain of threonine in the P3 position of SUMO-2 is oriented away from the surface of SENP1, and there is no clearly defined pocket that could be responsible for the tight binding of any amino



Figure 3 Subsite preferences of SENP1, 2, 5, 6, 7 and 8

SENPs were incubated with the PS-SCL in the presence or absence of 0.8 M sodium citrate, or C-terminally truncated SUMOs ( $\Delta$ ). The enzyme concentration was in the range 1–5  $\mu$ M. ACC production was monitored using an fMax multi-well fluorescence plate reader at an excitation wavelength of 355 nm and an emission wavelength of 460 nm with an assay time of 30 min. The *x*-axis represents the one-letter code of natural amino acids (0, norleucine). The *y*-axis represents the average relative activity expressed as a percentage of the best amino acid.



Figure 4 SENP binding modes

Comparison of the binding modes of the C-terminal residues of SUMOs or Nedd8 in the active sites of SENPs made on the basis of the available complex structure of these proteins from PDB accession numbers 2CKH, 1TGZ and 1XT9. Amino acid residue numbering is based on the assumption that the C-terminal glycine of mature SUMOs or Nedd8 is Gly-1.

acids. In contrast to the P3 position, the glutamine side-chain in P4 is oriented towards the surface of SENP1 and is located in a deep pocket formed mainly by phenylalanine, histidine and threonine side-chains of the enzyme, explaining the high specificity in this position. The side-chain of glutamine in the P5 position, like P3, points away from the enzyme, suggesting that P5 would also not dominate specificity. Unfortunately there is no crystal structure available for SENP5, but analogy with the results obtained for SENP1 and 2 suggests a similar mode of binding of the substrate around the active centre with tight binding and specificity in the P4 position and much lower specificity for amino acids in the P3 position. In the case of Group II, both SENP6 and 7 gave totally unexpected hits from the combinatorial library approach. Assay conditions using sodium citrate and C-terminally truncated SUMO-2 were required to obtain sufficient activity for analysis of these SENPs, and in both cases we observe a strong preference for leucine in the P4 position and less for the structurally closelyrelated isoleucine and norleucine. Tolerance for the expected glutamine residue was markedly weaker, giving a minimal increase of fluorescence for SENP6 and practically none for SENP7 (Figure 3). Preferences at P3 were also unexpected and favour arginine and tryptophan (Figure 3), with a slight increase in tyrosine preference in the presence of sodium citrate



Figure 5 Activation of SENP6 and 7

Activity was screened using the Ac-LRGG-AFC substrate. The *y*-axis represents the activity as an average of three individual assays. The enzyme concentration was 4  $\mu$ M (based on total protein) with a substrate concentration of 100  $\mu$ M and an assay time of 20 min.

(Supplemental Figure 1). Significantly, this profile is much closer to the Nedd8 or ubiquitin C-terminal sequence (LRGG) than SUMOs (QTGG). Intrigued by the unexpected subsite preferences for SENP6 and 7 obtained from our combinatorial scanning approach, we tested the specificity of activating factors using Ac-LRGG-AFC as substrate. Whereas truncated SUMO-2 and -3 could substantially enhance SENP6 and 7 activity, neither truncated SUMO-1 nor truncated Nedd8 nor ubiquitin had this ability (Figure 5). Consequently, although the observed activation by truncated SUMOs suggests exosite interactions are important for activity, the active site specificity of Group II SENPs suggests alternative natural substrates. Importantly, SENP6 and 7 revealed no detectable endopeptidase activity on the precursors of ubiquitin, Nedd8 and ISG15 (results not shown).

Positional scanning of SENP8 was tested only in low salt Tris buffer because this SENP was not substantially activated by  $\Delta$ C6-truncated Nedd8 or sodium citrate [14]. In the case of P4, the most preferred residue is leucine, which is the natural P4 residue of Nedd8. All other amino acids are considerably weaker, with only norleucine showing even 10% of the rate of leucine. In stark contrast with SENP1, 2, 5 and 6, but similar to SENP7, SENP8 demonstrates specificity at the P3 position, with a preference for arginine, which is the natural residue in Nedd8. Interestingly, tryptophan is tolerated almost as well as arginine. Our results are a good match to the crystal structure of SENP8 bound to Nedd8 (Figure 4) [21]. In particular, leucine in the P4 position is oriented toward the surface of SENP8 and is located in a narrow hydrophobic pocket, which would also explain the tolerance for norleucine at this position. Arginine in P3 is oriented away from the surface of SENP8, but interacts strongly via solvent molecules with the SENP8 surface, and these interactions enhance specificity. Arginine and tryptophan are quite different in size and structure, but share an important common feature, namely a secondary amine in the  $\epsilon$  position. This amine would be a likely

candidate as the major factor responsible for the equal preference of arginine and tryptophan at P3.

#### Characterization and validation of optimal substrates

From the PS-SCL preferences we selected a few sequences for resynthesis as single substrates to validate the scanning approach and define the best substrate for SENP1, 2 and 8. In the case of SENP1 and 2, we chose QTGG and QYGG recognition elements, as these represented the strongest candidates. We also synthesized pentapeptide substrates based on the sequence of SUMO-1 (EQTGG) or SUMO-2 and -3 (QQTGG) to determine whether extending the chain influenced substrate catalysis. Table 1 shows the  $k_{cal}/K_m$  values obtained for SENP1 and 2 in low salt Tris buffer, 0.8 M citrate buffer or in the presence of truncated SUMOs.

In the case of SENP1, the  $k_{cat}/K_m$  values for the tetrapeptide substrates Ac-QTGG-AFC and Ac-QYGG-AFC are very close, with a slight preference for threonine-containing substrate in the presence of low salt, and tyrosine-containing substrate in the presence of citrate buffer. The  $k_{cat}/K_m$  values increased dramatically in sodium citrate buffer (about 23-fold) or after incubation of SENP1 with truncated SUMOs at a 1:5 molar ratio (11-14-fold). Control experiments demonstrated that the 1:5 enzyme:truncated SUMO ratio was optimal for enhancing activity (results not shown). SENP1 did not discriminate between SUMOs in terms of enhancement of activity. The enhancement was specific for truncated SUMOs, since truncated Nedd8 at the same concentrations did not cause activation, nor did fulllength SUMOs [14]. Because of the relatively low solubility of Ac-QTGG-AFC (<100  $\mu$ M depending on assay conditions) we were not able to saturate the substrate-velocity relationship, and could not therefore confirm that cleavage of the tetrapeptide substrates follows a classic Michaelis-Menten relationship. There is no reason to think that cleavage of tetrapeptide substrates by SENPs is unusual, just that even the best rates were still rather low compared with  $k_{\text{cat}}/K_{\text{m}}$  values in the order of  $10^5 - 10^6 \text{ M}^{-1} \cdot \text{s}^{-1}$ , typical of highly potent endopeptidases [24].

Extending the peptide chain at the N-terminus did not provide any benefit to the substrates, indeed there was a substantial decline in catalytic rates in sodium citrate and truncated SUMO conditions when the enzymes were tested on pentapeptide substrates (Table 1). This is probably due to a partial overlap between the pentapeptide substrate with the acetyl-protecting group and truncated SUMO, rendering substrate-enhanced activity unfavourable. We have also synthesized heptapeptide substrates based on the sequences of SUMO-1 (Ac-YQEQTGG-AFC) and SUMO-2 and -3 (Ac-FQQQTGG-AFC), but these substrates are difficult to work with because of a tendency to precipitate and form gels in the assay, even at relatively low concentrations (low micromolar). Values for  $k_{cat}/K_m$  were hard to reproduce from assay to assay and therefore are omitted in this report.

Kinetic analysis of SENP2 on the various substrates produced results comparable with SENP1. The  $k_{cat}/K_m$  values in the case of low salt buffer are the lowest among all the conditions investigated. The tetrapeptide QTGG is minimally better compared with QYGG. In the case of sodium citrate buffer conditions, we observed a dramatic increase in the activity of SENP2. The  $k_{cat}/K_m$  values for both substrates are similar and the magnitude of increase of activity is around 33 and 40 times more for QTGG and QYGG substrates respectively. Truncated SUMO-2 and -3 produced activity enhancements similar to 0.8 M sodium citrate but, in stark contrast with SENP1, truncated SUMO-1 failed to enhance activity on any peptide substrate (Table 1). Similarly to SENP1, the incorporation of a fifth amino acid residue

#### Table 1 Kinetic analysis of preferred peptide substrates of human SENP1, 2 and 8

Data represent the mean  $\pm$  S.D. of three or more experiments. n.a., no significant activation by C-terminally truncated SUMO-1, -2, or -3 ( $\Delta$ C). Substrate concentration, 1–50  $\mu$ M; final enzyme concentration, 0.5–4  $\mu$ M.

		Tris buffer	$k_{\rm cal}/K_{\rm m}~({\rm M}^{-1}\cdot{\rm s}^{-1})$			
	Substrate		Sodium citrate	SUMO-1∆C	SUMO-2∆C	SUMO-3∆C
SENP1	Ac-QTGG-AFC Ac-QYGG-AFC Ac-EQTGG-AFC	13.97 (±1.6) 12.9 (±1.3) 8.97 (±1.4)	325.7 (±21.4) 302.7 (±30.7) 193.3 (±28.3)	191.3 (±11.3) 193.4 (±23.9) 34.5 (±5.1)	178.2 (±14) 175.4 (±16.7) 17.9 (±1.3)	148.9 (±44) 150.1 (±29.5) 17.3 (±0.5)
SENP2	Ac-QQTGG-AFC Ac-QTGG-AFC Ac-QYGG-AFC Ac-EQTGG-AFC	$20.2 (\pm 1.2) 6 (\pm 0.6) 4.7 (\pm 0.4) 1.77 (+ 0.2)$	286.8 (±22.3) 196.1 (±34.4) 189.8 (±27.3) 93.8 (±6.4)	35.2 ( <u>+</u> 3.9) n.a. n.a. n.a.	34.9 (±4.2) 202.6 (±25.1) 227.2 (±30.1) 14.5 (+2.0)	$\begin{array}{c} 34.4 (\pm 1.7) \\ 179 (\pm 9.2) \\ 203.3 (\pm 4.6) \\ 13.9 (\pm 1.0) \end{array}$
SENP8	Ac-QQTGG-AFC Ac-LRGG-AFC Ac-LWGG-AFC Ac-ALRGG-AFC	$5.5 (\pm 0.8) 7.5 (\pm 1.1) 5.7 (\pm 0.6) 5.9 (\pm 0.5)$	152.6 (±23)	n.a.	16.9 ( <u>+</u> 1.2)	19.9 ( <u>+</u> 1.0)

#### Table 2 Comparison of cleavage selectivity for SENP6, 7 and 8

Data represent the mean  $\pm$  S.D. of three or more experiments. Substrate concentration, 100–200  $\mu$ M; final enzyme concentration, 2.5–5  $\mu$ M. Units are nM AFC/ $\mu$ M enzyme/min. SENP6 and 7 were assayed in the presence of 0.8 M sodium citrate.

	SENP6	SENP7	SENP8
Ac-LRGG-AFC Ac-QTGG-AFC LRGG/QTGG ratio	$\begin{array}{c} 35.67 \pm 7.57 \\ 0.24 \pm 0.1 \\ 148 \end{array}$	$\begin{array}{c} 13.57 \pm 0.51 \\ 0.12 \pm 0.01 \\ 113 \end{array}$	$115.77 \pm 2.75 \\ 2.7 \pm 0.1 \\ 43$

to the substrate did not change the efficiency of cleavage by SENP2. The substrate sequence based on SUMO-1 (EQTGG) is the worst substrate of all those tested for SENP2 and around five times weaker than that based on SUMO-2 and -3 (QQTGG).

For SENP8, the  $k_{cat}/K_m$  values of the best substrates (LRGG and LWGG) are close to each other with a slight preference for LRGG, confirming observations from the PS-SCL. The pentapeptide substrate (ALRGG) conforming to Nedd8 (the natural substrate of SENP8) was no better than the corresponding tetrapeptide, indicating that extending the chain did not enhance catalysis.

The unexpected specificity of SENP6 and 7, revealed by the PS-SCL library, was confirmed by measuring cleavage efficiency of individual Ac-QTGG-AFC and Ac-LRGG-AFC substrates. The low catalytic efficiency of SENP6 and 7 made  $k_{cat}/K_m$  determinations difficult, so we report the rates in terms of relative specific activities (Table 2). Interestingly, the degree of preference for LRGG over QTGG was of the same magnitude for SENP6, 7 and 8.

#### P3 and P4 specificity in the context of full-length SUMO

There is a formal possibility that the P3 and P4 selectivity of SENPs observed using tetrapeptide sequences may not represent selectivity in the natural full-length (SUMO) substrates. For example, the conformational flexibility of the isolated tetrapeptides may be significantly larger than in the context of the full-length domains. To address this issue, we generated mutants of proSUMO-1 where the P3 position was varied. P3 positions that were predicted to be optimal from the PS-SCL (threonine, tyrosine and isoleucine) were cleaved with approximately equal efficiency by SENP1, but a poorly tolerated residue (glycine) was cleaved much less efficiently (Figure 6). To determine whether the unexpected P3 and P4 specificity of SENP6 and 7 would be revealed in the context of fulllength proSUMO, we compared the cleavage of full-length proSUMO-2 with a construct in which the P3 and P4 residues had been mutated to leucine and arginine. The proSUMO-2 proteins migrated differently in SDS/PAGE, although the cleaved products had an equivalent mobility, suggesting that the Cterminal tails may interact differentially with the glutaminethreonine versus leucine-arginine residues at P4 and P3 (Figure 7). Nevertheless, by using a range of SENP concentrations, we found that the leucine-arginine mutants are cleaved by a lower concentration of SENP6 and 7 than the wild-type sequence (Figure 7), revealing that the mutants act as better substrates than the wild-type in the context of full-length proSUMO-2. In contrast, the wild-type proSUMO-2 protein was substantially preferred to the leucine-arginine mutant by SENP1 (Supplemental Figure 2, http://www.BiochemJ.org/bj/409/bj4090461add.htm), confirming the predicted order of specificity applies in the context of full-length proSUMO substrates. These preferences support the results obtained with tetrapeptides, and confirms that SENP6 and 7 possess an unexpected P3 and P4 specificity.

#### DISCUSSION

In this report, we have defined the shortest peptidyl fluorogenic substrates that can be effectively cleaved by SENP1, 2, 5, 6, 7 and 8, and using a PS-SCL we have determined the optimal cleavage motif for each enzyme. Perhaps not too surprisingly, the optimal substrate for SENP1, 2 and 5 is the tetrapeptide sequence OTGG, identical to the natural substrates of these enzymes, namely SUMO-1, -2 and -3. Similarly, in the case of SENP8, the best sequence is the LRGG contained in its putative natural substrate Nedd8. The most surprising result was obtained for the Group II SENPs, suggesting that both SENP6 and 7 prefer the LRGG sequence, which corresponds more to Nedd8 or ubiquitinlike proteases. The results obtained from the PS-SCL reveal very high specificity of all tested SENPs in the P4 position and, in the case of SENP6, 7 and 8, substantial preferences also for the P3 position. SENP1, 2 and 5 reveal much less stringent requirements in the P3 position, tolerating a variety of side-chains, though with some preferences for bulky and branched side-chains.

In contrast to SENP8, the activity on synthetic substrates of the other SENPs tested here can be significantly modulated by buffer conditions. In all cases, the worst activity was observed in low salt Tris buffer, whereas sodium citrate, at high concentration, always increased the enzyme activity. Sodium citrate at 0.8 M is a Hofmeister-effect salt, with properties of ordering protein



Figure 6 Activity of SENP1 on wild-type and mutant proSUMO-1

Different amounts of SENP1 (2.3 to 200 nM, 1.5-fold dilution series) were incubated with the indicated proSUMO-1 (5  $\mu$ M) in a 30  $\mu$ I reaction mixture [50 mM Tris (pH 8.0), 20 mM NaCl and 5 mM DTT] for 1 h at 37 °C. The wild-type P3 threonine residue (panel A) was replaced by tyrosine (panel B), isoleucine (Panel C), or glycine (panel D). After incubation, the cleavage products were visualized by SDS/PAGE and stained with GelCode Blue. The grey triangle represents SENP concentration, and EC<sub>50</sub> represents the concentration of SENP1 required to process one-half of the proSUMO-1. S\* indicates substrate-only control (no SENP1).



Figure 7 Substrate specificity of SENP6 and 7

Analysis of SENP6 (**A**) and SENP7 (**B**) preference for Leu-Arg over Glu-Thr at the P4 and P3 positions using a full-length proSUMO protein substrate. Wild-type (QTGG) or mutant (LRGG) proSUMO-2 (5  $\mu$ M) was incubated with SENPs (62 nM to 1  $\mu$ M, 2 fold dilution series) in a 20  $\mu$ l reaction mixture for 1 h at 37 °C. After incubation, the cleavage products were visualized by SDS/PAGE and stained with GelCode Blue. The grey triangle represents SENP concentration, and EC<sub>50</sub> represents the concentration of SENP required to process one-half of the proSUMO-2. S<sup>\*</sup> indicates substrate-only control (no SENP).

structures, and it is likely that this effect causes a conformational change similar to C-terminally truncated SUMO, since the magnitudes of activation were very similar. The rather minimal changes in the preferences in the P3 and P4 positions seen with PS-SCL in the presence of sodium citrate or truncated SUMO suggest that any conformational changes around the active centre leading to this activation are not necessarily associated with substantial alterations in the specificity sites. A possible explanation of enhancement of activity comes from examining the crystal structures of the SUMOs bound to SENP2, where two main interaction regions between enzyme and full-length substrate can be distinguished [20,22]. The first interaction is around the active centre of the enzyme and involves four residues from the SUMO C-terminus, and the second one is a distance from the active centre, encompassing contact sites between the SUMO domain and the enzyme surface. We hypothesize that both the Hofmeister-effect and truncated SUMOs influence conformational changes only around the second recognition region between the proteins, facilitating access of the synthetic substrate to the active centre.

Overall, the  $k_{cat}/K_m$  values for the peptidyl substrates were substantially lower than for natural ones. For example, reported  $k_{cat}/K_{m}$  values for SENP2 on the natural substrate proSUMO are in the range of  $10^4$  to  $10^5$  M<sup>-1</sup> · s<sup>-1</sup> [25], which is up to two orders of magnitude higher than for our synthetic substrates. We propose that the lower turnover values for the synthetic substrates is probably due to an effect on  $K_{\rm m}$  rather than  $k_{\rm cat}$ . The natural and synthetic peptidyl substrates probably adopt an identical conformation with respect to occupancy of the active site cleft, as revealed by the comparison of selected mutants of full-length proSUMOs, and one would imagine that they undergo scission at similar catalytic rates  $(k_{cat})$ . But the substrates differ outside of the catalytic cleft, and therefore the binding step  $(K_m)$  would be very different. Our proposal is supported by the observation that pro-SUMO-1 and -2 substrates (which contain the same QTGG C-terminal motif, but different sequences within their respective SUMO domains) show comparable  $k_{cat}$  values, but the  $K_{\rm m}$  values differ by an order of magnitude [25]. Clearly the SUMO domain itself acts as an important exosite to decrease  $K_m$ , and our comprehensive analysis of synthetic substrates demonstrates the magnitude gain in catalytic efficiency afforded by the SUMO domain.

We observed no significant substrate activation in the case of SENP8, a property that distinguishes this SENP from all the other tested SENPs. Our analysis using a combinatorial scanning approach let us differentiate the individual SENPs by substrate preferences and mode of activation, and confirm that these enzymes form three specificity groups. Group I represents half of the family, consisting of SENP1, 2 and 5. These are typical deSUMOylating enzymes, and we expect the same for SENP3 (not tested here, but earlier reports reveal SUMO specificity) given its relatively close homology with SENP5, and previous reports demonstrating activity of the full-length SUMO precursor [9]. Group II members SENP6 and 7 are also activated by SUMO, but have a surprisingly distinct, and totally unpredicted, specificity in the active site cleft, preferring residues typical of Nedd8 or ubiquitin in the P3 and P4 positions. Group III consists of SENP8, and has specificity consistent with its previously reported activity on Nedd8 [6].

One of the objectives of this study was to determine whether small peptidyl substrates could be used to investigate specificity and activity of SENPs. Clearly they can, although catalytic rates are substantially lower than full-length SUMO or Nedd8 precursors. This is probably because the interactions between SENPs and the SUMO or Nedd8 domain enhance catalysis, as also revealed by our demonstration of SUMO-enhanced activity. The finding that SENP6 and 7 have an unexpected specificity suggests that SUMOs may not be their physiological substrates, and that the name deSUMOylating enzyme might be reconsidered in the nomenclature of these proteases. Neither of these enzymes was able to cleave precursors of Nedd8, ISG15, or ubiquitin, which contain the LRGG sequence at their C-terminal motif. We screened the full-length substrates corresponding with ubiquitin, ISG15 and Nedd8 containing the AMC fluorophore (Boston Biochem), without detecting cleavage by SENP6 or 7. We also observed no processing of ubiquitin-GFP (green fluorescent protein), or natural full-length proforms of Nedd8 and ISG15. by either SENP6 or 7, and we observed no processing of tetraubiquitin (Boston Biochem) chains by either SENP6 or 7 (results not shown). Possibly the natural substrate of SENP6 and 7 could be another protein with a ubiquitin-like fold, containing the LRGG-like sequence at its C-terminus, although our analysis of likely ubiquitin-like proteins in the literature and in protein sequence databases revealed no clues. Alternatively, the natural substrate may be a SUMO-conjugated species as suggested previously [8], although no obvious explanation comes to mind as to why the primary substrate specificity recognition elements in the active site cleft should be so different from the canonical deSUMOvlating enzymes SENP1, 2, 3 and 5.

Looking to the future, our demonstration of specific tetrapeptides as SENP substrates points to a route for the design of low-molecular-mass inhibitors or activity-based probes that are proving to be extremely useful in defining the functional roles of other types of cysteine proteases such as caspases or cathepsins [26].

We thank Prof. Jon Ellman (Department of Chemistry, University of California, Berkeley, CA, U.S.A.) for the gift of the ACC fluorophore and Prof. Keith Wilkinson (Department of Biochemistry, Emory University School of Medicine, Atlanta, GA, U.S.A.) and Miklos Bekes from our laboratory for stimulating discussions, and Scott Snipas for outstanding technical assistance. Supported in part by NIH grant U01 Al061139.

Received 17 July 2007/25 September 2007; accepted 5 October 2007 Published as BJ Immediate Publication 5 October 2007, doi:10.1042/BJ20070940

#### REFERENCES

- Melchior, F., Schergaut, M. and Pichler, A. (2003) SUMO: ligases, isopeptidases and nuclear pores. Trends Biochem. Sci. 28, 612–618
- 2 Hay, R. T. (2005) SUMO: a history of modification. Mol Cell. 18, 1-12
- 3 Yeh, E. T., Gong, L. and Kamitani, T. (2000) Ubiquitin-like proteins: new wines in new bottles. Gene. 248, 1–14
- 4 Hay, R. T. (2001) Protein modification by SUMO. Trends Biochem. Sci. 26, 332–333
- 5 Gill, G. (2004) SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? Genes Dev. 18, 2046–2059
- 6 Gan-Erdene, T., Nagamalleswari, K., Yin, L., Wu, K., Pan, Z. Q. and Wilkinson, K. D. (2003) Identification and characterization of DEN1, a deneddylase of the ULP family. J. Biol. Chem. 278, 28892–28900
- 7 Johnson, E. S. (2004) Protein modification by SUMO. Annu. Rev. Biochem. 73, 355–382
- 8 Mukhopadhyay, D., Ayaydin, F., Kolli, N., Tan, S. H., Anan, T., Kametaka, A., Azuma, Y., Wilkinson, K. D. and Dasso, M. (2006) SUSP1 antagonizes formation of highly SUM02/3-conjugated species. J. Cell Biol. **174**, 939–949
- 9 Gong, L. and Yeh, E. T. (2006) Characterization of a family of nucleolar SUMO-specific proteases with preference for SUMO-2 or SUMO-3. J. Biol. Chem. 281, 15869–15877
- 10 Zhang, H., Saitoh, H. and Matunis, M. J. (2002) Enzymes of the SUMO modification pathway localize to filaments of the nuclear pore complex. Mol. Cell Biol. 22, 6498–6508
- 11 Gong, L., Millas, S., Maul, G. G. and Yeh, E. T. (2000) Differential regulation of sentrinized proteins by a novel sentrin-specific protease. J. Biol. Chem. 275, 3355–3359
- 12 Dang, L. C., Melandri, F. D. and Stein, R. L. (1998) Kinetic and mechanistic studies on the hydrolysis of ubiquitin C-terminal 7-amido-4-methylcoumarin by deubiquitinating enzymes. Biochemistry **37**, 1868–1879
- 13 Borodovsky, A., Ovaa, H., Meester, W. J., Venanzi, E. S., Bogyo, M. S., Hekking, B. G., Ploegh, H. L., Kessler, B. M. and Overkleeft, H. S. (2005) Small-molecule inhibitors and probes for ubiquitin- and ubiquitin-like-specific proteases. Chembiochem. 6, 287–291
- 14 Mikolajczyk, J., Drag, M., Bekes, M., Cao, J. T., Ronai, Z. and Salvesen, G. S. (2007) Small ubiquitin-related modifier (SUMO)-specific proteases: profiling the specificities and activities of human SENPs. J. Biol. Chem. 282, 26217–26224
- Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins. Biochemistry 6, 1948–1954
- 16 Alves, L. C., Almeida, P. C., Franzoni, L., Juliano, L. and Juliano, M. A. (1996) Synthesis of Nα-protected aminoacyl 7-amino-4-methyl-coumarin amide by phosphorous oxychloride and preparation of specific fluorogenic substrates for papain. Pept. Res. 9, 92–96
- 17 Backes, B. J. and Ellman, J. A. (1999) An alkanesulfonamide "safety-catch" linker for solid-phase synthesis. J. Org. Chem. 64, 2322–2330
- 18 Maly, D. J., Leonetti, F., Backes, B. J., Dauber, D. S., Harris, J. L., Craik, C. S. and Ellman, J. A. (2002) Expedient solid-phase synthesis of fluorogenic protease substrates using the 7-amino-4-carbarnoylmethylcoumarin (ACC) fluorophore. J Org. Chem. 67, 910–915
- 19 Choe, Y., Leonetti, F., Greenbaum, D. C., Lecaille, F., Bogyo, M., Bromme, D., Ellman, J. A. and Craik, C. S. (2006) Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. J. Biol. Chem. 281, 12824–12832
- 20 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
- 21 Reverter, D., Wu, K., Erdene, T. G., Pan, Z. Q., Wilkinson, K. D. and Lima, C. D. (2005) Structure of a complex between Nedd8 and the Ulp/Senp protease family member Den1. J. Mol. Biol. 345, 141–151
- 22 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 paralogues during processing. Biochem. J. **397**, 279–288
- 23 Harris, J. L., Backes, B. J., Leonetti, F., Mahrus, S., Ellman, J. A. and Craik, C. S. (2000) Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. Proc. Natl. Acad. Sci. U.S.A. 97, 7754–7759
- 24 Thornberry, N. A., Rano, T. A., Peterson, E. P., Rasper, D. M., Timkey, T., Garcia-Calvo, M., Houtzager, V. M., Nordstrom, P. A., Roy, S., Vaillancourt, J. P. et al. (1997) A combinatorial approach defines specificities of members of the caspase family and granzyme B. Functional relationships established for key mediators of apoptosis. J. Biol. Chem. **272**, 17907–17911
- 25 Reverter, D. and Lima, C. D. (2006) Structural basis for SENP2 protease interactions with SUMO precursors and conjugated substrates. Nat. Struct. Mol. Biol. 13, 1060–1068
- 26 Kato, D., Boatright, K. M., Berger, A. B., Nazif, T., Blum, G., Ryan, C., Chehade, K. A., Salvesen, G. S. and Bogyo, M. (2005) Activity-based probes that target diverse cysteine protease families. Nat. Chem. Biol. 1, 33–38