



LETTER

An improved method for identifying SUMOylation sites of viral proteins

Dear Editor,

In recent years, post-translational modifications (PTMs) by small ubiquitin-related modifiers (SUMOs) have emerged as an important regulatory mechanism for both cellular and viral processes (Ribet and Cossart, 2010). Identifying the SUMOylation sites of the target protein is important to understand the molecular mechanism underlying SUMO modification and virus-host interactions, as well as provide new insights into antiviral drug development (Wimmer and Schreiner, 2015). Traditional site-directed mutagenesis for identifying viral protein SUMOylation sites lacks a specific aim and is laborious (McManus et al., 2016). Recently, mass spectrometry (MS) has been employed as an accurate and sensitive tool to identify PTM sites, thereby greatly expanding the number of known SUMOylated proteins (Pedrioli et al., 2006). However, during viral infection, SUMOylation is highly dynamic and SUMOylated viral proteins often have low abundance, which makes studying SUMOylation under natural conditions difficult.

The eukaryotic cell expression system has been used to generate endogenously modified proteins, but with low yields of SUMOylated proteins, and further antibody-dependent problems in purification hampered the sensitivity during discovery of viral SUMOylated proteins (Rosas-Acosta et al., 2005; Hsiao et al., 2012). Since SUMOylation is mainly based on a three-enzyme reaction (SUMO-activating enzyme, E1; SUMO-conjugating enzyme, Ubc9; and E3 ligase), a cell-free system with these proteins was tested and found to be successful, as the influence of cellular SUMO-specific proteases (SENPs) and other SUMOylated proteins was absent (Knuesel et al., 2005). Later, an *Escherichia coli* SUMOylation system consisting of integrated His-tagged SUMO-1, E1, and Ubc9 in one plasmid and a GST-tagged target protein in another plasmid was reported; this system could be used to harvest a large amount of SUMOylated proteins by protein purification from a bacterial strain within 2 days (Weber et al., 2014). To enable mass spectrometry detection, the length of SUMO modified peptide should be limited, as the long side chain on the target lysine residues resulting in highly complicated fragmentation

patterns in MS/MS analyses, which makes interpretation of data difficult (Matic et al., 2008; Blomster et al., 2010). SUMO-1, a modified SUMO molecule with a threonine 95 to arginine (T95R) mutation, was developed to overcome this problem. The diglycine (GG) tag produced by trypsin digestion of SUMO-1 (T95R) could represent both SUMOylation sites and be easily identified by classical MS detection (Knuesel et al., 2005; Impens et al., 2014).

Considering all the above known information, we developed an improved method to accurately detect SUMOylation sites of viral proteins, which combines an *E. coli* system carrying the components required for SUMOylation with modified SUMO-1 (T95R) and MS detection. Taking the viral protein enterovirus 71 (EV71) 3D polymerase as an example, our system provides an efficient and practical strategy for the identification of viral protein SUMOylation sites.

The experimental scheme of 3D protein detection is depicted in [Figure 1A](#). To generate SUMOylated GST-3D-SUMO-1 (T95R), pGEX-6p-3D plasmid was transformed into pSUMO-1 (T95R)-containing *E. coli* BL21 (DE3) cells; we created this *E. coli* strain for protein purification as previously described (Weber et al., 2014; Liu et al., 2016). However, we introduced several modifications to the previously published protocol—the cell lysates were first loaded on a GStrap HP column (GE Healthcare) and then on a HisTrap HP column (GE Healthcare) for enrichment ([Figure 1A](#)). The collected samples were analyzed using SDS-polyacrylamide gel (SDS-PAGE) electrophoresis followed by Coomassie brilliant blue staining ([Figure 1B](#)). After two separation steps, 3D-GST-SUMO-1 (T95R) was harvested, and the SUMOylated 3D showed the same band pattern as wild-type pSUMO-1 ([Figure 1C](#)).

For MS detection, SUMOylated 3D bands above 100 kDa were excised from the gel and subjected to in-gel digestion (Matic et al., 2010; Impens et al., 2014) ([Figure 1C](#)). The trypsin-digested peptides were purified, and then detected via reversed-phase liquid chromatography-electrospray ionization-mass spectrometer (RPLC-ESI-MS)/MS. Liquid chromatography-MS (LC-MS)/MS detection was carried out on a hybrid quadrupole-TOF LC/MS/MS

mass spectrometer (TripleTOF 5600 +, AB Sciex) equipped with a nanospray source. Raw data from TripleTOF 5600 + were analyzed using ProteinPilot Software. MS analysis showed that K159, K24, and K127 were modified by SUMO-1, and modified K159 was present at a higher frequency in the peptides with scores higher than those of the two other sites. This result was in accordance with that of our previous study, which showed that K159 is the major SUMOylation site for SUMO-1 and SUMO-3 (Liu et al., 2016). The peptides identified by MS are displayed in [Supplementary Figure S1](#). The 3D

protein possesses 27 lysine residues, but using this bacterial system MS analysis, we only detected 3 lysine residues, indicating that this system was unbiased and specific lysine residues were preferentially modified via SUMOylation in the *E. coli* system.

To fully evaluate this system, the lysine residues identified by MS were mutated into arginine residues and verified using the cell lysates for immunoprecipitation (IP) experiments. K24R, K127R, and double mutants of K24R/K159R and K127R/K159R were created to compare with previously identified K159R, K159R/150–152

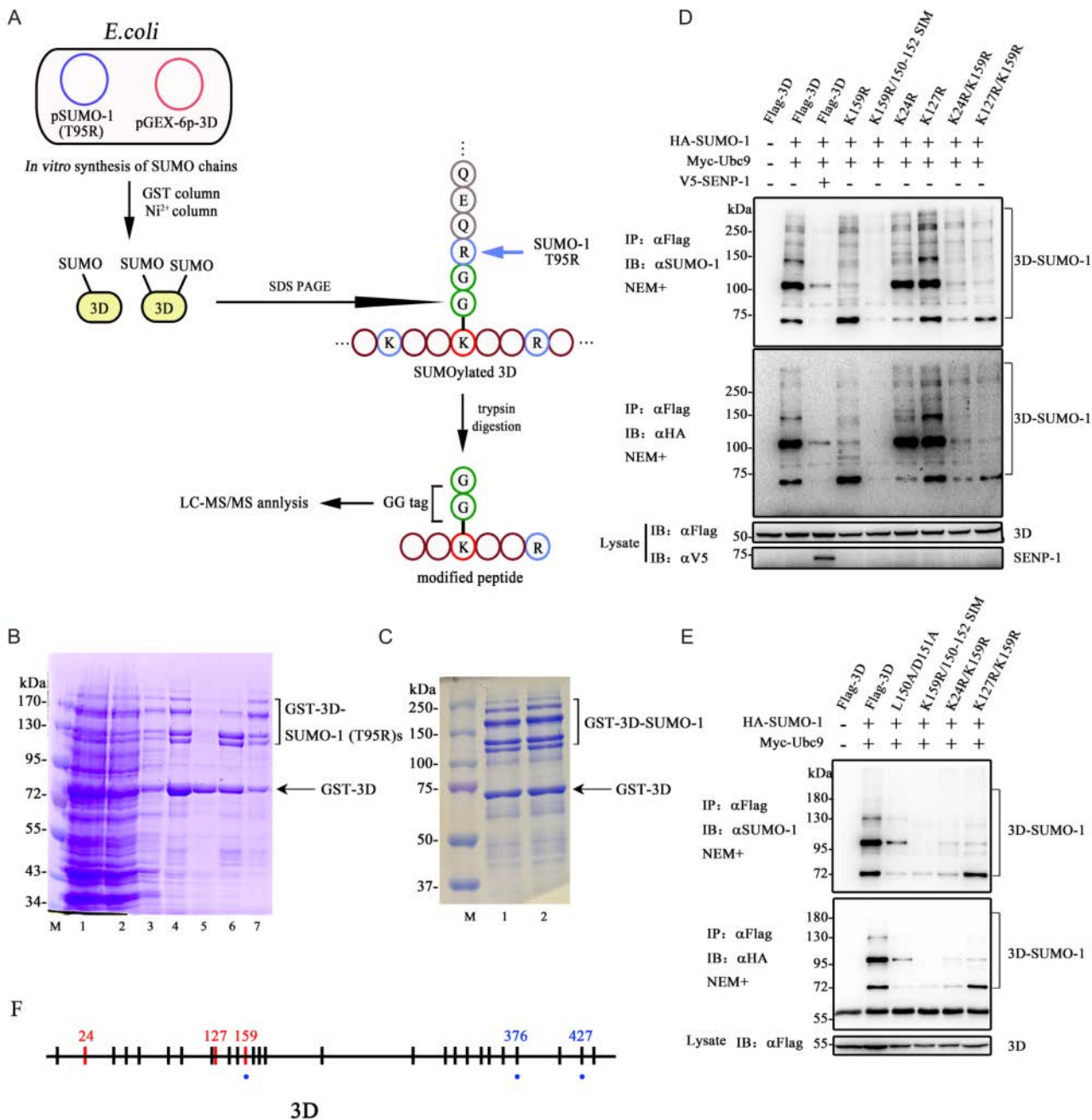


Figure 1. Identification of SUMOylation sites of enterovirus 71 3D protein by mass spectrometry (MS) in *E. coli* system. (A) Flowchart of the method used to identify SUMOylation sites in this study. The modified SUMOylated 3D protein with SUMO-1 (T95R) could produce a diglycine (GG) tag in the resulting tail after digested by trypsin. Residues are depicted in circles. The SUMO target lysine is depicted in a red circle, and the diglycine (GG) tag is depicted in green circles. (B) SDS-PAGE analysis of SUMO-1 (T95R) modified 3D protein expressed in *E. coli* system. M, marker. Lanes 1–7: crude lysate supernatants, GST column flow through, GST column-eluted fraction by PBS, dialyzed elution fractions by glutathione, nickel column flow through, 10% imidazole-eluted fraction, and 30% imidazole-eluted fraction. (C) The comparison of SUMOylated 3D expressed by the pSUMO-1 and pSUMO-1 (T95R) plasmid in *E. coli* system after two separation steps. M, marker; Lane 1, sample purified through GST column and nickel column from bacteria bearing pSUMO-1; Lane 2, sample purified through GST column and nickel column from bacteria bearing pSUMO-1 (T95R). (D, E) Verification of newly identified SUMOylation sites by immunoprecipitation and immunoblot analyses. K24R, K127R, and double mutants of K24R/K159R and K127R/K159R were created to compare with previously identified K159R, K159R/150–152 SIM, and L150A/D151A. 3D and its mutants were co-expressed with Ubc9 and SUMO-1 in human embryonic kidney cells (HEK 293T) to perform SUMOylation assays. V5-SEN1 was introduced for deSUMOylation. Co-IP assay using anti-FLAG antibody and the western blot was carried out with anti-HA/ anti-SUMO-1 antibody. (F) The diagram of lysine residues in 3D amino acid sequence. The sites identified by MS analysis were colored in red, which were K24, K127, and K159. The top three lysine residues from online prediction (www.abgent.com/sumoplot) were dotted with blue underneath, which were K159, K376, and K427.

SIM, and L150A/D151A. 3D and its mutants were co-expressed with Ubc9 and SUMO-1 in human embryonic kidney cells (HEK 293T) to perform SUMOylation assays. V5-SEN1 was introduced for deSUMOylation. Figure 1D shows that K127R exerted no apparent effect on SUMOylation, whereas the bands modified by SUMO-1 (around 70 kDa) were considerably decreased in K24R. Double-mutant K24R/K159R decreased the majority of the SUMOylated bands (Figure 1D, 1E). K159R/150–152 SIM and L150A/D151A were the promising mutants that remarkably reduced the modified proteins (Figure 1D, 1E). These results indicate that K159 and K24 are the major SUMOylation sites and 150–152 SIM is also involved in 3D SUMOylation. Therefore, this system not only confirmed the known SUMOylation sites of EV71 3D, but also expanded our understanding about 3D SUMOylation sites.

Typically, lysine residues responsible for SUMO modification are within a canonical consensus motif (Matic *et al.*, 2010). The top three lysine residues identified by online SUMOylation prediction based on this consensus motif were K159, K376, and K427 (www.abgent.com/sumoplot), but only K159 was proved to be valid (Liu *et al.*, 2016) (Figure 1F, Supplementary Figure S2). Moreover, K24 does not follow the consensus motif; therefore, bioinformatics prediction has missed this site, and to identify sites similar to this, we must resort to MS analysis for accurate identification. The identification of K127 may be caused by false positives, which may be due to non-specific SUMOylated bands in the bacterial system or the limitation of MS. The SUMOylation study of menin, a tumor regulator, showed that this protein may bear multiple SUMOylation sites; however, MS only detected lysine 591 as a major modification site (Feng *et al.*, 2013). Despite the result indicating that MS presents false positive

sites or shortage in finding multiple SUMOylation sites, this method still shows a strong advantage compared with the time-consuming mutation of lysine residues or when the amount of SUMOylated protein is low similar to that of a viral protein. Since MS only detected 3 lysine residues in 3D, IP experiment showed that 2 lysine residues were responsible for 3D SUMOylation, and K159 presents the highest frequency. This system could mimic the natural status of a protein modified by SUMO-1 in cells to some extent.

Our system showed enhanced convenience and efficiency over the other reported *E. coli* system. Compared with the *E. coli* system applied in the study on AtMYB30, a transcription factor of *Arabidopsis*, which requires three plasmids for SUMOylation and was more customized for the study on plant SUMOylation sites, our system needed only two plasmids for SUMOylated protein (Okada *et al.*, 2009; Weber *et al.*, 2014). Compared with the study on SUMOylation sites of UL44 of human cytomegalovirus, our bacterial system was improved by Schar *et al.* to enhance efficiency and specificity (Sinigalia *et al.*, 2012; Weber *et al.*, 2014). Moreover, the system in UL44 is highly reliable on the use of database “ChopNSpice” to overcome limitation of the long length of SUMO-1 after trypsin digestion. Our method of applying mutant SUMO-1 (T95R) to increase the coverage of SUMOylation sites during MS analysis is more reliable and efficient. (Sinigalia *et al.*, 2012).

In this study, we improved the most commonly used methods for MS detection of SUMOylated sites and used an *E. coli* recombination SUMOylation system with SUMO-1 (T95R). This system provides fast enrichment of SUMOylated viral protein in less than 2 days, and shows advantage over the method of collecting modified protein from cells in convenience and sensitivity.

Furthermore, this method provides an option for rapid and accurate identification of the potential viral protein SUMOylation sites.

FOOTNOTES

This work was supported by the National Natural Science Foundation of China [grant number 81471953]; and the Youth Innovation Promotion Association of CAS [grant number 2016302]. We thank the Core Facility and Technical Support, Wuhan Institute of Virology for their help. This article does not contain any studies with human or animal subjects performed by any of the authors. The authors declare that they have no conflict of interest.

Supplementary figures are available on the websites of *Virologica Sinica*: www.virosin.org; link.springer.com/journal/12250.

Yan Liu¹, Zhongyuan Tan¹, Bo Shu¹, Yuan Zhang¹,
Caishang Zheng¹, Xianliang Ke¹, Xi Chen²,
Hanzhong Wang¹, Zhenhua Zheng¹✉

1. CAS Key Laboratory of Special Pathogens and Biosafety,
Center for Emerging Infectious Diseases, Wuhan Institute of
Virology, Chinese Academy of Sciences, Wuhan 430071, China

2. Wuhan Institutes of Biotechnology, Wuhan 430075, China

✉Correspondence:

Phone: +86-27-87197295, Fax: +86-27-87199239,

Email: zhengzh@wh.iov.cn

ORCID: 0000-0002-4117-9364

Published online: 31 August 2017

REFERENCES

- Blomster HA, Imanishi SY, Siimes J, et al. 2010. *J Biol Chem*, 285: 19324–19329.
- Feng ZJ, Gurung B, Jin GH, et al. 2013. *Am J Cancer Res*, 3: 96–106.
- Hsiao HH, Meulmeester E, Urlaub H. 2012. *Methods Mol Biol*, 893: 431–441.
- Impens F, Radoshevich L, Cossart P, et al. 2014. *Proc Natl Acad Sci U S A*, 111: 12432–12437.
- Knuesel M, Cheung HT, Hamady M, et al. 2005. *Mol Cell Proteomics*, 4: 1626–1636.
- Liu Y, Zheng Z, Shu B, et al. 2016. *J Virol*, 90: 10472–10485.
- Matic I, Schimmel J, Hendriks IA, et al. 2010. *Mol Cell*, 39: 641–652.
- Matic I, van Hagen M, Schimmel J, et al. 2008. *Mol Cell Proteomics*, 7: 132–144.
- McManus FP, Altamirano CD, Thibault P. 2016. *Nat Protoc*, 11: 387–397.
- Okada S, Nagabuchi M, Takamura Y, et al. 2009. *Plant Cell Physiol*, 50: 1049–1061.
- Pedrioli PG, Raught B, Zhang XD, et al. 2006. *Nat Methods*, 3: 533–539.
- Ribet D, Cossart P. 2010. *Cell*, 143: 694–702.
- Rosas-Acosta G, Russell WK, Deyrieux A, et al. 2005. *Mol Cell Proteomics*, 4: 56–72.
- Sinigalia E, Alvisi G, Segre CV, et al. 2012. *Plos One*, 7.
- Weber AR, Schuermann D, Schar P. 2014. *PLoS One*, 9: e102157.
- Wimmer P, Schreiner S. 2015. *Viruses*, 7: 4854–4872.