Published in final edited form as:

Bioorg Med Chem Lett. 2015 ; 25(10): 2146–2151. doi:10.1016/j.bmcl.2015.03.069.

Synthetic derivatives of the SUMO consensus sequence provide a basis for improved substrate recognition

Melissa J. Leyva^a, Yeong Sang Kim^a, Megan L. Peach^b, and John S. Schneekloth Jr.^{a,*}

^aChemical Biology Laboratory, National Cancer Institute, 376 Boyles St., Frederick, MD 21702, USA

^bBasic Science Program, Chemical Biology Laboratory, Leidos Biomedical Research, Inc., National Cancer Institute, 376 Boyles St., Frederick, MD 21702, USA

Abstract

Protein sumoylation is a dynamic posttranslational modification that regulates a diverse subset of the proteome. The mechanism by which sumoylation enzymes recognize their cognate substrates is unclear, and the consequences of sumoylation remain difficult to predict. While small molecule probes of the sumoylation process could be valuable for understanding SUMO biology, few small molecules that modulate this process exist. Here, we report the synthesis and evaluation of over 600 oxime-containing peptide sumoylation substrates. Our work demonstrates that higher modification efficiency can be achieved with non-natural side chains that deviate substantially from the consensus site requirement of a hydrophobic substituent. Furthermore, docking studies suggest that these improved substrates mimic binding interactions that are used by other endogenous protein sequences through tertiary interactions. The development of these high efficiency substrates provides key mechanistic insights toward specific recognition of low molecular weight species in the sumoylation pathway.

Keywords

Combinatorial chemistry; Conjugation; Enzymes; Peptide mimics; SUMO

The posttranslational modification of protein substrates with the small ubiquitin-like modifier (SUMO) is a highly dynamic process critical to many cellular phenomena. Similar to ubiquitin, SUMO utilizes a cascade of E1, E2, and E3 enzymes to catalyze covalent conjugation to protein substrates.^{1–3} In addition to being required for cell cycle progression and development,⁴ DNA damage repair,⁵ survival of heat shock,⁶ and protection from ischemia,⁷ this pathway is dysregulated in a number of cancers.^{8–10} A growing number of proteins within the genome have been identified as substrates for SUMO-1, -2, and/or -3 modification,^{11,12} and the consequences of this modification are currently an area of significant investigation. Although great strides have been made in structural biology, the

*Corresponding author. Tel.: +1 301 228 4620; fax: +1 301 846 6033. schneeklothjs@mail.nih.gov (J.S. Schneekloth).

Supplementary data

Supplementary data (compound synthesis and characterization, assay protocols, and modeling figures) associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bmcl.2015.03.069>.

specific mechanism by which so many structurally diverse substrates within the proteome are recognized by sumoylation enzymes remains in question. The biological significance of this conjugation pathway emphasizes the need for new methods to interrogate molecular recognition events of sumoylating enzymes.

One unique aspect that differentiates the sumoylation pathway from other ubiquitin-like signaling pathways is that Ubc9, the SUMO E2 enzyme, directly recognizes a substrate consensus sequence.¹³ This sequence is typically described as ψ KxE, where ψ is a hydrophobic amino acid, K is the lysine where the SUMO isopeptide bond is formed, x is any amino acid, and E is an acidic amino acid. This sequence is found in a broad variety of proteins, such as I κ B α , p53, PML, and AdE1B, where it typically occurs in exposed flexible loops. In addition to this canonical sequence, several variant consensus sequences have been found, including inverted sequences,⁴ hydrophobic clusters,¹⁴ negatively charged consensus motifs,¹⁵ and phospho-dependent sumoylation sequences.¹⁶ These short sequences provide at least one basis for specificity in substrate recognition in the sumoylation pathway. We reasoned that development of a small, more efficient substrate might provide insight into the molecular basis of substrate recognition by Ubc9, or serve as a starting point for the identification of small molecule ligands for the enzyme. To the best of our knowledge, these efforts are the first attempts to derivatize the sumoylation consensus sequence in order to identify substrates with improved efficiency, though one synthetic substrate has been serendipitously identified via a screening approach.¹⁷

Here, we report the identification of small, synthetic substrates that undergo SUMO modification with superior efficiency than endogenous peptide sequences. We utilize an oxime scanning approach to prepare and evaluate 600 chemically diverse oxime-containing non-natural peptide sequences as substrates for sumoylation (Fig. 1). Two of these peptides were found to be significantly more reactive than the native peptide sequence itself, and were studied in further detail. In silico modeling experiments provide a rationale for the improved recognition of these substrates relative to natural sequences.

We previously reported the use of a 10-mer peptide derived from the androgen receptor and labeled with the fluorophore 5-FAM (FL-AR) as a competent substrate in a microfluidic electrophoretic mobility shift biochemical assay.¹⁸ It is well known that certain short peptide sequences derived from endogenous substrates will recognize Ubc9 directly and undergo sumoylation in biochemical assays, even in the absence of E3 ligases.^{19,20} Generally, these peptides must contain the ψ KxE sequence, and substantial deviations from this sequence decrease modification efficiency. To investigate the ability of shorter sequences to undergo sumoylation, we evaluated a series of four AR-derived peptides as competitive substrates in our biochemical assay. This assay utilizes a microfluidic system to separate fluorescently labeled substrate and product of an enzymatic reaction on the basis of differing electrophoretic mobility. Under optimized separation conditions, both the substrate FL-AR peptide and sumoylated product could be observed. Thus, percent conversion could be quantified ratiometrically. In this study, we measured an IC₅₀ for each peptide for inhibition of FL-AR sumoylation in order to compare substrate efficiencies (Fig. 2A). Thus, each peptide was evaluated for its ability to outcompete sumoylation of the fluorescent FL-AR peptide substrate. In this assay the AR-derived 10-mer, 8-mer, 6-mer, and 4-mer peptides

were active as competitive substrates, with IC_{50} values of 9 ± 0.1 , 6 ± 0.2 , 8 ± 0.3 , and $30 \pm 7 \mu M$, respectively. It is notable that the simple tetrapeptide, composed of the consensus sequence alone, was competent in the assay, albeit slightly less so than the longer sequences.

Having established that the AR-derived tetrapeptide is a competent substrate for SUMO modification, we evaluated five other tetrapeptide sequences from known sumoylation substrates. We synthesized and evaluated sequences from RanGap1, $I\kappa B\alpha$, p53, PML, and AdE1B as competitive substrates in dose response (Fig. 2B).¹³ Here, Ac-IKLE-NH₂ (the sequence from AR) and Ac-IKME-NH₂ (the sequence from PML) were the most active peptides, with an IC_{50} of $30 \pm 7 \mu M$ and $33 \pm 2 \mu M$, respectively. On the basis of these experiments, we elected to pursue Ac-IKLE-NH₂ as the scaffold on which to synthesize our library.

In order to rapidly investigate large numbers of modified peptides related to the SUMO consensus sequence as competitive substrates, we utilized an oxime scanning approach. Oxime bond ligation has shown to be a useful strategy to rapidly generate a library of compounds.^{21–30} Some of the advantages of using oxime ligation over other click reactions are that it can be conducted using commercially available aldehydes and reaction products can be directly biologically evaluated without purification. Six different peptides were prepared by solid phase peptide synthesis, each of which contained an aminooxy (AO) group replacing a different residue in the consensus sequence or at the C-terminus of the peptide (Fig. 3A). Each peptide was then condensed with a chemically diverse library of 100 aldehydes to generate a library of 600 SUMO consensus sequence mimetics. Each oxime-containing peptide was then individually evaluated at a concentration of 25 μM in our previously reported biochemical sumoylation assay.

The results of this screen revealed that substrate efficiency depends to a striking extent on the identity of the modified residue. The vast majority of the 600 assayed peptides were substantially less active than the consensus sequence-containing peptide itself, highlighting the specificity required for Ubc9 to recognize its substrates (Fig. S1 in Supporting information). Only 10 substrates derivatized at the ψ - and C-terminal positions displayed comparable or improved activity relative to Ac-IKLE-NH₂. These compounds were further evaluated after resynthesis and purification but only a few gave reproducible results. From this analysis, we identified compounds **3** and **4** as the most active peptides (Fig. 3B and C). Furthermore, we validated that activity was not due to unreacted aldehyde and a significant difference in percent conversion was not observed between purified compounds **3** and **4** and their crude reaction mixtures (Fig. S2). Purified peptides **3** and **4** were at least two-fold more potent competitive substrates than Ac-IKLE-NH₂ with IC_{50} values of $12 \pm 0.8 \mu M$ and $15 \pm 1 \mu M$, respectively. Furthermore, the activities of **3** and **4** are comparable to the full-length AR 10-mer peptide itself, but with only four residues. The unmodified aminooxy derivative Ac-AO-KLE-NH₂ was a competent but three-fold worse substrate than Ac-IKLE-NH₂, confirming that both the pyrazole group in peptide **4** and the pyridone group in peptide **3** confer a pronounced improvement in the ability of each peptide to serve as a competitive sumoylation substrate. It is of note that the canonical consensus sequence requires a hydrophobic residue in this position whereas both **3** and **4** are relatively polar heterocycles with multiple hydrogen bond donors/acceptors.

Next, a small series of focused analogs of both **3** and **4** were prepared and evaluated as competitive substrates to investigate their structure–activity relationship (Fig. 4). For oxime **3**, simple pyridone (**5**), cyanopyridine (**6**), aminopyridine (**7**), and pyridine (**8**) derivatives all retained superior activity to Ac-IKLE-NH₂. However, a phenyl derivative (**9**) exhibited somewhat decreased activity, implying that lack of the pyridine nitrogen may be associated with poor activity. In parallel, several derivatives of oxime **4** were prepared. Pyrazole derivatives substituted in the 3-position with methyl (**10**), methoxy (**11**), or other amino groups (**12** and **13**) retained activity. However, an *N*-cyclopentyl pyrrole analog (**14**) displayed no inhibitory activity, indicating that similar to the pyridone series, deviation from the core heterocycle is detrimental.

In order to confirm that the synthetic peptides were acting as substrates and not causing the observed effects by a non-specific mechanism, gel shift assays were performed using a Western blot. Peptides **3** and **4** were exposed to purified recombinant E1, E2, and SUMO-1 proteins in the presence of ATP. After quenching the assay with EDTA, samples were analyzed by Western blot using an anti-SUMO-1 antibody (Figs. 5 and S3). The formation of the SUMO–peptide conjugate was observed in a time-dependent fashion, confirming that the synthetic peptides are in fact being modified by SUMO-1.

Next, we pursued modeling studies to probe potential structural features of Ubc9 that might be responsible for the increased activity of peptides **3** and **4**. We employed the covalent docking mechanism in GOLD using the highest resolution structure of Ubc9 available (PDB: 2GRN),³¹ in which the central core of the peptide is treated as part of the receptor to fix the position of the lysine and glutamate residues, while optimizing the position and receptor interactions of the ψ -residue and C-terminal oxime-linked fragments.

We found that peptides **3** and **4** dock most favorably in extended structures, where the oxime-linked fragment stretches along the Ubc9 surface (Fig. S4). The predicted binding location is a shallow groove on the Ubc9 surface, adjacent to the active site.³² The predicted interactions made by peptides **3** and **4** with Ubc9 include van der Waals interactions with residues Pro 128, Ala 131, and Tyr 134, as well as a hydrogen bond between Thr 135 and the oxygen in the morpholine ring of peptide **4** (Fig. 6). Notably RanGAP1, the substrate originally characterized in complex with Ubc9 (PDB: 2GRN) utilizes these same key binding interactions by bringing a neighboring helix into proximity of the consensus sequence (Fig. S5). It appears that these short synthetic peptides may be able to mimic tertiary interactions made by larger endogenous sumoylation substrates.

We have used an oxime scanning approach to synthesize and screen 600 oxime-based synthetic derivatives of the sumoylation consensus sequence. This approach led to the identification of two classes of synthetic substrates that undergo sumoylation with greater efficiency than that of endogenous peptide sequences. As expected, derivatives that replace the critical lysine or acidic residues in the sequence show no activity. However, replacing the hydrophobic ψ residue of the consensus sequence with either of two specific polar heterocyclic substituents resulted in substantially improved activity. Although natural amino acid mutagenesis studies have suggested that hydrophobic residues are required for efficient sumoylation, this study demonstrates that non-natural polar residues can provide

improvements. Subsequent biochemical analysis confirmed that the synthetic peptides discovered here act as substrates and that SUMO conjugation could be observed in vitro. Finally, modeling studies provide insight into how these peptides may recognize Ubc9 and gain greater efficiency as substrates. This work indicates that rather than utilizing hydrophobic interactions, the pyridone and pyrazole substituents likely take advantage of aromatic stacking interactions with Y134 and hydrogen bonding interactions with T135 of Ubc9. Thus, our work demonstrates that replacing natural side chains with specific non-natural side chains can provide substrates that recognize Ubc9 through alternative recognition mechanisms. This mode of interaction may be similar to how RanGap1 recognizes Ubc9. RanGap1 utilizes tertiary structure to achieve higher affinity by positioning a helix within proximity of the consensus site and making contacts with Y134 and T135 (among other residues) on Ubc9.

In addition to providing knowledge about how Ubc9 recognizes substrates, the identification of a low molecular weight sumoylation substrate may eventually prove to be a starting point for the development of a substrate competitive, low molecular weight chemical inhibitor. While chemical inhibition of the sumoylation pathway has been reported in the literature,^{33–37} it has been challenging to identify inhibitors through traditional high throughput screening approaches.¹⁷ Future efforts will be aimed toward leveraging these newly identified extended binding motifs toward continued improvement of the affinity of both peptides **3** and **4** for Ubc9 or development of a peptidomimetic inhibitor of Ubc9.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Intramural Research Program of the National Institutes of Health, Center for Cancer Research, and the National Cancer Institute, National Institutes of Health. This work was also funded in part with Federal funds from the National Cancer Institute, National Institutes of Health, under contract HHSN261200800001E. The authors thank Dr. Terrence R. Burke, Jr. (NCI, NIH) for providing the aldehydes for the oxime ligation reactions. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organizations imply endorsement by the U.S. Government.

References and notes

1. Gareau JR; Lima CD Nat. Rev. Mol. Cell Biol 2010, 11, 861. [PubMed: 21102611]
2. Geiss-Friedlander R; Melchior F Nat. Rev. Mol. Cell Biol 2007, 8, 947. [PubMed: 18000527]
3. Johnson ES Annu. Rev. Biochem 2004, 73, 355. [PubMed: 15189146]
4. Nacerddine K; Lehembre F; Bhaumik M; Artus J; Cohen-Tannoudji M; Babinet C; Pandolfi PP; Dejean A Dev. Cell 2005, 9, 769. [PubMed: 16326389]
5. Psakhye I; Jentsch S Cell 2012, 151, 807. [PubMed: 23122649]
6. Golebiowski F; Matic I; Tatham MH; Cole C; Yin Y; Nakamura A; Cox J; Barton GJ; Mann M; Hay RT Sci. Signal 2009, 2, ra24. [PubMed: 19471022]
7. Lee YJ; Mou Y; Maric D; Klimanis D; Auh S; Hallenbeck JM Plos One 2011, 6, e25852. [PubMed: 22016779]

8. Driscoll JJ; Pelluru D; Lefkimiatis K; Fulciniti M; Prabhala RH; Greipp PR; Barlogie B; Tai YT; Anderson KC; Shaughnessy JD, Jr.; Annunziata CM; Munshi NC Blood 2010, 115, 2827. [PubMed: 19965618]
9. Driscoll JJ; DeChowdhury R Target Oncol. 2010, 5, 281. [PubMed: 21125340]
10. Kessler JD; Kahle KT; Sun T; Meerbrey KL; Schlabach MR; Schmitt EM; Skinner SO; Xu Q; Li MZ; Hartman ZC; Rao M; Yu P; Dominguez-Vidana R; Liang AC; Solimini NL; Bernardi RJ; Yu B; Hsu T; Golding I; Luo J; Osborne CK; Creighton CJ; Hilsenbeck SG; Schiff R; Shaw CA; Elledge SJ; Westbrook TF Science 2012, 335, 348. [PubMed: 22157079]
11. Tatham MH; Matic I; Mann M; Hay RT Sci. Signal 2011, 4, rs4. [PubMed: 21693764]
12. Vertegaal AC; Andersen JS; Ogg SC; Hay RT; Mann M; Lamond AI Mol. Cell. Proteom. MCP 2006, 5, 2298.
13. Rodriguez MS; Dargemont C; Hay RT J. Biol. Chem 2001, 276, 12654. [PubMed: 11124955]
14. Matic I; Schimmel J; Hendriks IA; van Santen MA; van de Rijke F; van Dam F; Gnäd F; Mann M; Vertegaal AC Mol. Cell 2010, 39, 641. [PubMed: 20797634]
15. Yang SH; Galanis A; Witty J; Sharrocks AD EMBO J. 2006, 25, 5083. [PubMed: 17036045]
16. Hietakangas V; Anckar J; Blomster HA; Fujimoto M; Palvimo J; Nakai A; Sistonen L Proc. Natl. Acad. Sci. U.S.A 2006, 103, 45. [PubMed: 16371476]
17. Brandt M; Szewczuk LM; Zhang H; Hong X; McCormick PM; Lewis TS; Graham TI; Hung ST; Harper-Jones AD; Kerrigan JJ; Wang DY; Dul E; Hou WF; Ho TF; Meek TD; Cheung MH; Johanson KO; Jones CS; Schwartz B; Kumar S; Oliff AI; Kirkpatrick RB Assay Drug Dev. Technol 2013, 11, 308. [PubMed: 23772552]
18. Kim YS; Nagy K; Keyser S; Schneekloth JS Chem. Biol 2013, 20, 604. [PubMed: 23601649]
19. Lin D; Tatham MH; Yu B; Kim S; Hay RT; Chen YJ Biol. Chem 2002, 277, 21740.
20. Yunus AA; Lima CD Methods Enzymol. 2005, 398, 74. [PubMed: 16275321]
21. Maly DJ; Choong IC; Ellman JA Proc. Natl. Acad. Sci. U.S.A 2000, 97, 2419. [PubMed: 10716979]
22. Jiang YL; Krosky DJ; Seiple L; Stivers JT J. Am. Chem. Soc 2005, 127, 17412. [PubMed: 16332091]
23. Johnson SM; Petrassi HM; Palaninathan SK; Mohamedmohaideen NN; Purkey HE; Nichols C; Chiang KP; Walkup T; Sacchettini JC; Sharpless BK; Kelly JW J. Med. Chem 2005, 48, 1576. [PubMed: 15743199]
24. Chung S; Parker JB; Bianchet M; Amzel LM; Stivers JT Nat. Chem. Biol 2009, 5, 407. [PubMed: 19396178]
25. Bahta M; Liu F; Kim SE; Stephen AG; Fisher RJ; Burke TR, Jr. Nat. Protoc 2012, 7, 686. [PubMed: 22422315]
26. Liu F; Park JE; Qian WJ; Lim D; Scharow A; Berg T; Yaffe MB; Lee KS, ; Burke TR, Jr. ACS Chem. Biol 2012, 7, 805. [PubMed: 22292814]
27. Bahta M; Burke TR, Jr. ChemMedChem 2011, 6, 1363. [PubMed: 21671403]
28. Bahta M; Lountos GT; Dyas B; Kim SE; Ulrich RG; Waugh DS; Burke TR, Jr. J. Med. Chem 2011, 54, 2933. [PubMed: 21443195]
29. Liu F; Hakami RM; Dyas B; Bahta M; Lountos GT; Waugh DS; Ulrich RG, ; Burke TR, Jr. Bioorg. Med. Chem. Lett 2010, 20, 2813. [PubMed: 20350805]
30. Liu F; Stephen AG; Waheed AA; Aman MJ; Freed EO; Fisher RJ; Burke TR, Jr. ChemBioChem 2008, 9, 2000. [PubMed: 18655064]
31. Verdonk ML; Cole JC; Hartshorn MJ; Murray CW; Taylor RD Proteins-Struct. Funct. Genet 2003, 52, 609. [PubMed: 12910460]
32. Bernier-Villamor V; Sampson DA; Matunis MJ; Lima CD Cell 2002, 108, 345. [PubMed: 11853669]
33. Fukuda I; Ito A; Hirai G; Nishimura S; Kawasaki H; Saitoh H; Kimura K; Sodeoka M; Yoshida M Chem. Biol 2009, 16, 133. [PubMed: 19246003]
34. Fukuda I; Ito A; Uramoto M; Saitoh H; Kawasaki H; Osada H; Yoshida MJ Antibiot. 2009, 62, 221.

35. Hirohama M; Kumar A; Fukuda I; Matsuoka S; Igarashi Y; Saitoh H; Takagi M; Shin-ya K; Honda K; Kondoh Y; Saito T; Nakao Y; Osada H; Zhang KYJ; Yoshida M; Ito A ACS Chem. Biol 2013, 8, 2635. [PubMed: 24143955]
36. Takemoto M; Kawamura Y; Hirohama M; Yamaguchi Y; Handa H; Saitoh H; Nakao Y; Kawada M; Khalid K; Koshino H; Kimura K; Ito A; Yoshida MJ Antibiot. 2009, 62, 221.
37. Kumar A; Ito DA; Hirohama M; Yoshida M; Zhang KY J. J. Chem. Inf. Model 2014, 54, 2784.

Figure 1.
Schematic representation of an oxime scanning strategy to identify highly active sumoylation consensus peptide mimics.

Figure 2.

Assessment of synthetic polypeptides as competitive sumoylation substrates in a microfluidic electrophoretic mobility shift biochemical assay. Effects of (A) varied peptide length and (B) varied sequence (six consensus sequences from endogenous sumoylation substrates). All peptides were prepared by solid phase peptide synthesis, are acetylated at the N-terminus, and amidated at the C-terminus. Error bars represent 1 standard deviation from the mean ($n = 3$). When error bars are not visible, they are smaller than the size of the symbols. Percent conversion is normalized to DMSO control.

Figure 3.

An oxime-based diversification approach to identifying more efficient peptide sumoylation substrates. (A) Strategy for preparation of the library of oxime-containing peptides. (B) Structures and (C) potency of the three most active substrates identified from the screen. Error bars represent 1 standard deviation from the mean ($n = 3$). When error bars are not visible, they are smaller than the size of the symbols. Percent conversion is normalized to DMSO control.

Figure 4. Structure–activity relationship of oxime-containing peptide substrates. Numbers in parenthesis represent percent conversion relative to a DMSO control, with errors reported as the one standard deviation of the mean ($n = 3$).

Figure 5. Time course experiments demonstrating formation of SUMO-1-peptide adducts by Western blot using an anti-SUMO-1 antibody.

Figure 6. Docked structures of (A) peptide 4 and (B) peptide 3 with Ubc9. Ubc9 is in grey, peptides are in green, and oxime side chains are in cyan and yellow. Cys93 is the catalytic cysteine.