J. Biochem. **144,** 39–49 (2008)
doi:10.1093/jb/mvn039

SUMO Assay with Peptide Arrays on Solid Support: Insights into SUMO Target Sites

Klaus Schwamborn[1,*], Puck Knipscheer[2,†], Evert van Dijk[1], Willem J. van Dijk[2], Titia K. Sixma[2], Rob H. Meloen[1,3] and Johannes P.M. Langedijk[1]

[1]Pepscan Therapeutics BV, Zuidersluisweg 2, 8243 RC Lelystad; [2]Department of Molecular Carcinogenesis, the Netherlands Cancer Institute and Center for Biomedical Genetics, Plesmanlaan, Amsterdam; and [3]Academic Biomedical Centre, University of Utrecht, Yalelaan 1, Utrecht, the Netherlands

Received February 11, 2008; accepted March 10, 2008; published online March 15, 2008

The modification of proteins by SUMO (small ubiquitin-like modifier) regulates various cellular processes. Sumoylation often occurs on a specific lysine residue within the consensus motif ψKxE/D. However, little is known about the specificity and selectivity of SUMO target sites. We describe here a SUMO assay with peptide array on solid support for the simultaneous characterization of hundreds of different SUMO target sites. This approach was used to characterize known SUMO substrates. The position of the motif within the peptide and the amino acids flanking the acceptor site affected the efficiency of SUMO modification. Interestingly, a sequence of only four amino acids, corresponding to the SUMO consensus motif without flanking amino acids, was a *bona fide* target site. Analysis of a peptide library for all variants of the ψKxE/D consensus motif revealed that the first and third positions in the tetrapeptide preferably contain aromatic amino acid residues. Furthermore, by adding the SUMO E3 ligase PIAS1 to the reaction mixture, we show specific enhancement of the modification of a PIAS1-dependent SUMO substrate in this system. Overall, our results demonstrate that the sumoylation assay with peptide array on solid support can be used for the high-throughput characterization of SUMO target sites, and provide new insights into the composition, selectivity and specificity of SUMO target sites.

Key words: assay, E3 ligase, peptide library, pepscan, SUMO (small ubiquitin-like modifier).

Post-translational modification by small ubiquitin-like modifiers (SUMO) is an important mechanism regulating a diversity of cellular functions, including sub-cellular distribution, gene transcription and protein stability (*1–3*). Sumoylation has also been shown to be linked with the pathogenesis of various disorders, including Alzheimer's and Huntington's disease (*4*, *5*), highlighting the importance of this protein modification. It remains unclear how many SUMO substrates exist in the human proteome; however, recent proteomic screening approaches have identified new targets and the number of validated substrates is rapidly increasing (*6–10*).

SUMO proteins are structurally related to ubiquitin and are expressed as precursors that undergo proteolytic cleavage, resulting in the availability of the C-terminal glycine–glycine motif for conjugation. At least four mammalian SUMO proteins—SUMO-1, -2, -3 and -4—have been identified. SUMO-2 and -3 have amino acid sequences 96% identical to each other and about 50% identical to that of SUMO-1. SUMO-4 is more similar to SUMO-2 and -3 than to SUMO-1. SUMO is covalently conjugated by an isopeptide bond to lysine residues in the substrates. Sumoylation is a three-step process, similar to ubiquitination, involving an E1-activating enzyme (Aos1/Uba2 heterodimer), an E2-conjugating enzyme (Ubc9) and E3 ligases, including Ran-binding protein-2, Polycomb-2 and PIAS proteins (*1–3*, *11–13*). However, recombinant E1, Ubc9, and SUMO are sufficient for the ATP-dependent SUMO modification of substrates *in vitro*. Consistent with this observation, biochemical and structural studies have shown that the SUMO consensus motif ψKxE/D (where ψ is a hydrophobic amino acid, x is any amino acid and K is the site of SUMO conjugation) is directly recognized by the Ubc9 active site (*14–17*). Several groups have suggested that there may be an extended SUMO consensus motif. For example, the synergy control motif (SC) is defined by the presence of proline residues flanking the core SUMO motif (*18*). An extended sumoylation motif—NDSM (negatively charged amino acid-dependent sumoylation motif)—has also been identified in several SUMO substrates (*19*). Moreover, clusters of acidic residues downstream from the core motif were also found in the transcription factor Elk-1 and the CRD1 domain of p300 (*20*). In line with the acidic stretch, bioinformatics analyses recently identified a subset of SUMO consensus sites called PDMS (phosphorylation-dependent sumoylation motifs) (*19*, *21*), combining a classical SUMO consensus site with an adjacent proline-directed phosphorylation site, ψKxExxSP, as in the major

*To whom correspondence should be addressed. Tel: +31 320 225300, Fax: +31 320 225301,
E-mail: k.schwamborn@pepscan.com
†Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.

MEF2 isoforms, for example. However, non-consensus sumoylation sites have also been reported for several SUMO targets (*2*, *22*).

The identification and validation of new SUMO substrates and deciphering of the SUMO acceptor site in a given protein are of considerable interest. Bioinformatics studies based on target-site search engines (*23*, *24*) and mass spectrometry of *in vitro* sumoylated proteins are very useful for this (*25*). A chip-based analysis of SUMO conjugation to a target protein has recently been developed (*26*). This report provided proof of principle that a sumoylation reaction can be performed on a solid support (glass slides). However, this study was based exclusively on the use of the RanGAP1 protein domain as a model SUMO substrate. Two other SUMO assays based on fluorescence resonance energy transfer (FRET) technology have been established: one evaluates SUMO protease activity (*27*) and the other monitors SUMO modification in solution in a high-throughput system (*28*). However, the two FRET-based assays have in common that the systematic investigation of SUMO substrates is possible only by testing one target at a time. Hence, there is currently no direct experimental tool for simultaneously characterizing large numbers of SUMO target sites in identical experimental conditions.

We describe here a sumoylation assay with peptide array on solid support. Using this assay format, we show that the amino acids flanking the consensus SUMO motif contribute to the selectivity and specificity of the modification. A known PIAS1 substrate could be validated by adding the recombinant SUMO E3 ligase to the reaction. We also defined, in detail, the amino acids present in each position of the consensus motif ψKxE/D. The high-throughput SUMO assay format with peptide array technology might also be useful for drug discovery and diagnosis.

MATERIALS AND METHODS

Reagents and Antibodies—Mouse α-SUMO1 (sc-5308), biotin-coupled mouse IgG and horseradish peroxidase (HRP)-conjugated streptavidin was obtained from Santa Cruz, DAKO and Southern Biotech, respectively. ABTS [2,2′-azino-bis(3-ethylbenzthiazoline-6-sulphonic acid)] was obtained from Roche. Recombinant PIAS1 was purchased from Biomol.

Peptide Synthesis—Peptides were synthesized, according to standard procedures, on an Applied Biosystems 430A synthesizer or on a Hamilton Microlab 2200 (Reno, NV), using Fmoc/HBTU (2′′-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) chemistry. The peptide preparations were more than 90% pure, as determined by analytical liquid chromatography/mass spectrometry.

Cyclic Peptides—Cyclic peptides were synthesized as previously described (*29*). Briefly, the reaction is an SN2-type reaction in which the thiol (RSH/RS-) group of a cysteine reacts as a nucleophile with the (activated) bromomethyl group of a benzylbromide derivative. It is based on the extremely rapid and efficient cyclization reaction between cysteine containing peptides α,α-dibromoxylene derivatives. The reaction runs smoothly at room temperature in <1 h without the need for any type of catalysis. The main advantage is that peptides can be used without any form of side chain protection, as a result of the exquisite selectivity of this reaction for thiols. The reaction is generally performed in aqueous solution (ACN/H_2O) at 0.5–1.0 mM peptide-concentration under slightly basic conditions (pH ∼7.8). The corresponding disulphide (RSSR), formed upon oxidation of the thiol (e.g. by O_2), is a potential by-product under the conditions used.

Pepscan Cards—Peptide synthesis on a solid support was performed as previously described (*30*, *31*).

Sumoylation Reaction with Peptides on a Solid Support—Standard conditions for the sumoylation reaction were as follows: 0.15 μM E1, 0.20 μM E2 and 0.4 μM SUMO-1. The reaction buffer contained 5 mM ATP, 50 mM NaCl, 5 mM $MgCl_2$, 0.2 M DTT, 1% BSA and 3% Tween-20. The peptide cards—peptides were presynthesized and chemically coupled to a solid support or directly synthesized on a solid support—were incubated with the reaction mixture for 30 min at 37°C. Nonspecifically bound proteins were then washed off by sonication for 2 min in a bath containing 1% SDS, 0.1% β-mercaptoethanol and 100 mM Na_2HPO_4. Further nonspecific binding was blocked by incubation with BSA/ps (4% BSA, 5% horse serum, 1% Tween-80 in PBS) for 30 min. We tested for SUMO conjugation by serial reactions with anti-SUMO-1 (0.2 μg/ml), secondary anti-mouse IgG-biotin (1/2,000), anti-streptavidin-HRP (1/2,000) in BSA/ps, with final detection in an ABTS-dependent enzyme reaction, according to the manufacturer's instructions. A modification efficiency-dependent colour reaction developed in the presence of SUMO conjugates. The optical density of the colour (green) was quantified using a CCD camera and an image processing system. The optical densities measured with the CCD camera (OD_{CCD}) are measured between 550 and –800 nm (orange filter). Using this filter, the green colour becomes a grey value. The set-up consists of a CCD camera and a 55-mm lens (Sony CCD Video Camera XC-77RR, Nikon micronikkor 55 mm f/2.8 lens), a camera adaptor (Sony Camera adaptor DC-77RR) and the image processing software package TIM, v. 3.36 (Difa Measuring Systems). The data presented are the means ± SD of at least two independent experiments.

Cell Culture and Total Cell Lysate Preparation—HeLa cells were maintained in DMEM medium (Invitrogen) supplemented with 10% fetal calf serum, 100 U/ml penicillin and 100 μg/ml streptomycin at 37°C, under an atmosphere containing 5% CO_2. Cells were lysed by adding an equal volume of 1 × CHRIS buffer (50 mM Tris, pH 8.0, 0.5% NP-40, 200 mM NaCl and 0.1 mM EDTA), supplemented with a protease inhibitor cocktail from Boehringer Manheim.

Protein Purification—E1 (Aos1/Uba2), Ubc9 and SUMO-1 were cloned as previously described (*12*, *22*, *32*). GST-RanGAP (420–589) was cloned in pETM30 (European Molecular Biology Laboratory). Proteins were purified by standard techniques and based on previously described protocols (*12*), briefly: unless otherwise stated, all protein purification protocols involved overnight IPTG-induced expression at 15°C in *Escherichia coli* Rosetta (DE3) or *E. coli* BL21(DE3) pLysS for

Uba2 (Novagen), bacterial lysis using sonication, and centrifugation at 18,000*g* for 1 h to collect soluble proteins. Lysis buffer contained 20 mM Tris pH 8.0, 40 to 200 mM NaCl, 5 mM β-mercaptoethanol, 1 mM PMSF and one tablet EDTA-free protease inhibitors (Roche) per 50 ml. After the specific purification steps described below, aliquots of the proteins were flash frozen, and stored at −80°C. His-Aos1 and Uba2 were separately expressed; cultures were mixed in a 10:1 ratio of Uba2:Aos1 expressing cells before the harvesting centrifugation step. The soluble protein fraction was applied to Ni-beads, eluted with imidazole. The eluate was applied to a MonoQ column and eluted with a salt gradient between 0.1 and 0.5 M NaCl. The final purification step was on a Superdex S200 gel filtration column in 20 mM Tris pH 8.0, 200 mM NaCl, 0.1 mM PMSF and 1 mM DTT. Ubc9 was purified on a POROS S column in 20 mM Bis–Tris pH 6.5, 50 mM NaCl, 0.1 mM PMSF, 1 mM DTT, and eluted with a salt gradient between 0.05 and 1 M NaCl. Ubc9 containing fractions were concentrated, pooled and further purified by gel filtration on a Superdex 75 column in 20 mM Tris pH 8.0, 100 mM NaCl, 0.1 mM PMSF and 1 mM DTT. Purification of SUMO-1 involved a 10% PEI (polyethylene imine) precipitation after which the soluble fraction was loaded on a Superdex200 26/60 gel filtration column in 20 mM Tris pH 8.8, 50 mM NaCl, 0.1 mM PMSF and 1 mM DTT. Finally, the protein was purified on a MonoQ column in the same buffer where it eluted at around 150 mM NaCl. GST-RanGAP (420–589) was purified using GST affinity purification followed by a gel filtration step on a Superdex200 in 20 mM Tris pH 8.0, 100 mM NaCl, 0.1 mM PMSF and 1 mM DTT.

RESULTS

SUMO Modification in Solution—In vitro assays usually involve the incubation of substrates (produced in bacteria) with proteins required for the SUMO reaction. E1 (Uba2/Aos1), E2 (Ubc9) and SUMO are often sufficient for the ATP-dependent SUMO modification of substrates in solution (*16*, *33*). Before investigating assay conditions directly on a solid support, the reactivity of all purified proteins was validated by conducting an *in vitro* sumoylation experiment in solution (Fig. 1). For this purpose, the Ran GTPase-activating protein RanGAP1 was selected, a well described model SUMO substrate and one of the most efficient SUMO targets identified to date (*12*). We prepared a construct encoding the RanGAP1 protein domain [(amino acids 420–589), encompassing the SUMO target site at position 526] fused to GST, with an expected molecular weight of about 50 kDa. Incubation of the RanGAP1 protein domain with all the relevant proteins for sumoylation, including E1, E2 and SUMO-1, resulted in a shift of the band to about 70 kDa, consistent with efficient modification of RanGAP1 (Fig. 1). The reaction conditions were based on previous reports (*14*, *26*, *28*, *34*, *35*); however, increasing the incubation time to 6 h, resulted in more efficient modification than observed with an incubation time of 1 h (data not shown). Thus, all the purified components required for SUMO modification were biologically active and could be used for the reaction on a solid support.

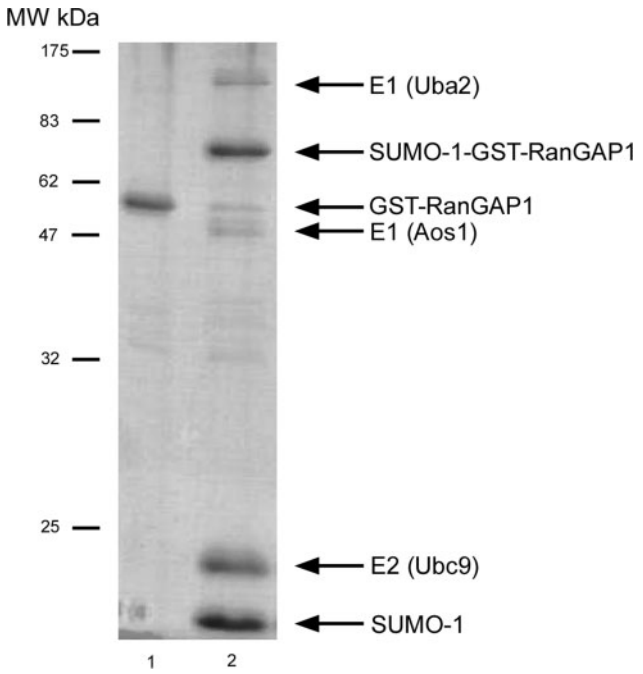

Fig. 1. **Sumoylation of RanGAP1 in solution.** The GST-fused RanGAP1 protein domain (0.2 mg/ml) was incubated with 0.69 μM E1, 1.5 μM E2 und 10 μM SUMO-1 for 6 h at 37°C. Proteins were resolved by SDS–PAGE in a 12.5% polyacrylamide gel, which was then stained with Coomassie blue. Unmodified and SUMO-modified RanGAP1, and the other proteins required for SUMO modification, are indicated. Molecular weight (MW) in kilo Dalton is also indicated.

SUMO Assay with Pre-Synthesized Peptides Chemically Coupled to a Solid Support—Small peptides, 11- to 15-amino residues long, are efficiently sumoylated in solution (*15*, *22*, *35*). The RanGAP1 protein domain has also been shown to function as a SUMO substrate on glass slides, providing the first evidence for *in vitro* sumoylation on a solid support (*26*). The goal was to combine these two approaches in the development of a SUMO assay with peptide arrays on a solid support. Unlike sumoylation in solution, the SUMO reaction on a solid support must be optimized in terms of both the composition of the reaction buffer and the concentrations of the relevant proteins. The chemical composition of the surface of the solid support may also be critical, as it may influence both the true sumoylation signal and false signals generated by non-specific protein binding.

The assay was optimized with a set of 10 RanGAP1 peptides, linked chemically via a cysteine residue to the surface of grafted polypropylene cards (*31*). The set included peptides of different length (9–15 residues), containing the SUMO target site LKSE. Control peptides including variants at the C-terminus were used in which the critical Lys in the consensus motif as well as the other downstream Lys was replaced by an Arg residue (Table 1). Three microlitres of SUMO reaction mixture was added to each well and incubated for 6 h at 37°C. Modified peptides were detected by incubation with anti-SUMO antibodies followed by an ABTS-dependent-enzyme linked immunoassay. The extent of sumoylation reaction was analysed

Table 1. **RanGAP1 peptide variants.**

Peptide number	Peptide sequence
1	**CGLLKSEDK**
2	CGLLRSEDK
3	CGLLRSEDR
4	**CHMGLLKSEDKVK**
5	CHMGLLRSEDKVK
6	CHMGLLRSEDRVK
7	CHMGLLRSEDRVR
8	CHMGLLRSEDKVR
9	**CVHMGLLKSEDKVKA**
10	CVHMGLLRSEDKVKA

Amino acid residues in the peptides carrying a Lys in the consensus motif are indicated in bold letters. The acceptor Lys and corresponding Arg in the control peptides are underlined.

with an image quantification program (see MATERIALS AND METHODS section). The RanGAP1 peptides, including the negative control peptides (compare peptide 1 with peptides 2 and 3, peptide 4 with peptides 5–8 and peptide 9 with peptide 10), were all sumoylated to similar extents, indicating that 'in-solution' conditions led to the non-specific binding of SUMO to the peptides or the card surface (Fig. 2A). We tried to inhibit non-specific protein binding by pre-treatment of the cards with the blocking reagent BSA, together with Tween-20. However, this did not result in a clear increase in signal-to-noise ratio (data not shown). In contrast, the addition of BSA and Tween-20 to the reaction mixture improved the specificity of the sumoylation reaction on the card (Fig. 2B). Next, the concentration of the E1, Ubc9 and SUMO components

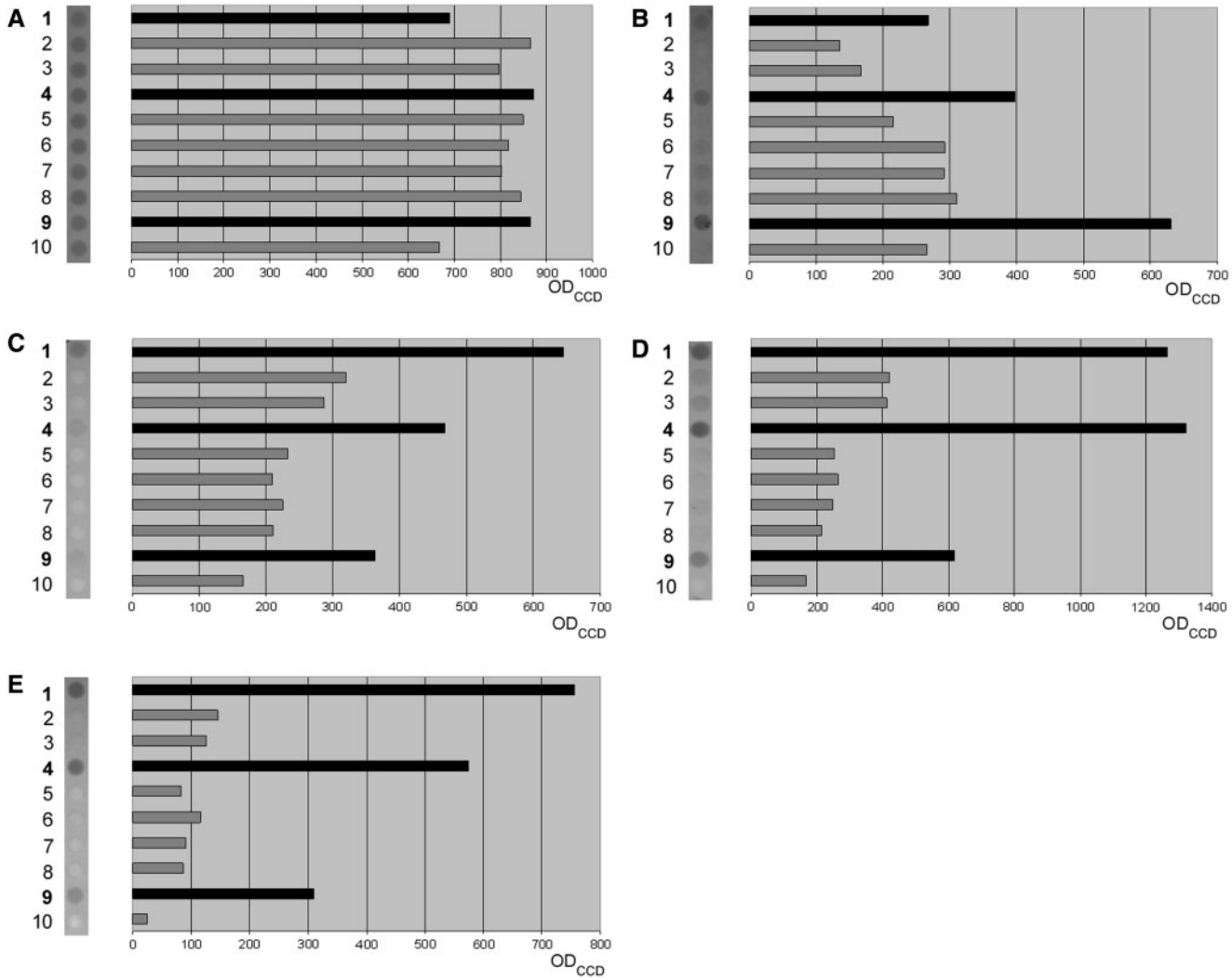

Fig. 2. **Assay optimization.** Sumoylation assay with pre-synthesized peptides chemically coupled to a solid support performed under the following conditions: (A) 0.69 μM E1, 1.5 μM E2 und 10 μM SUMO-1 for 6 h at 37°C. (B) As in (A) but in the presence of 1% BSA and 3% Tween-20. (C) Low concentration of each component: 0.15 μM E1, 0.20 μM E2 und 0.4 μM SUMO-1 for 6 h at 37°C. (D) As in (C), but incubation for 30 min. (E) As in (C), but before detection, modified peptide cards were incubated in a sonication bath for 2 min. Numbers correspond to RanGAP1 peptides in Table 1. On the left side of each panel, a picture of the investigated peptides after detection is shown. Each spot represents an individual SUMO reaction in one well. On the right side of each panel, a graph shows the corresponding OD_{CCD} values along the *x*-axis.

was adjusted; we found that, in general, lowering the concentration of all components increased the specificity of the SUMO reaction (Fig. 2C), indicating that a large excess of all factors resulted in unspecific binding. Decreasing the incubation time from 6 h to 30 min also increased the specificity of the reaction (Fig. 2D). Finally, we carried out a washing step directly after the SUMO reaction, to remove non-covalently conjugated SUMO. This procedure, involving sonication in a bath, further improved the performance of the assay (Fig. 2E).

Thus, we have developed a SUMO assay, involving pre-synthesized peptide arrays coupled to a solid support. Different RanGAP1 peptides displayed different intensities of sumoylation: The shortest wild-type RanGAP1 peptide, CGLLKSEDK, had the highest sumoylation value, at about 750, followed by CHMGLLKSEDKIK (~570) and the longest peptide CIHMGLLKSEDKIKA (~300) (Fig. 2E). However, the control peptides 2 and 3 of the shortest RanGAP1 version also gave high values, at about 120. Thus, the longest version (peptide 9) in relation to the control peptide 10 had the strongest SUMO-specific activity, by a factor of about 12.

SUMO Assay with Peptides Directly Synthesized on a Solid Support ('Pepscan Cards')—As shown before, sumoylation of pre-synthesized peptides coupled to a solid support is specific and efficient. We then investigated whether it was feasible to carry out the SUMO reaction with peptides directly synthesized on a solid support, as the chemistry of these types of peptide cards is different (*30*, *31*). Application of the sumoylation reaction on these cards is very important, because large peptide arrays can only be generated when peptides are synthesized directly on a solid support. Three RanGAP1 peptides with different length and corresponding control peptides (target Lys is replaced by Arg) were directly synthesized on cards, and the SUMO assay was then carried out in the conditions established above. We calculated the extent of SUMO modification for each reaction by dividing the value of the wild-type peptide with the value obtained for the control peptide after SUMO modification; this value we call the sumoylation factor, SF. The RanGAP1 peptides were all sumoylated, however the SF varied between 2 and 3 (Fig. 3), suggesting that the different peptide length influences the sumoylation efficiency. Although the SF of SUMO-modified peptides seemed to be lower (compared with the peptides in Fig. 2E), the experiment showed that Pepscan cards (peptide synthesis on cards) could be used for sumoylation assays on a solid support. Further efforts to improve the SUMO assay on Pepscan cards were not successful (data not shown), but we nonetheless continued our studies with this type of card, as its application is crucial for applying peptide array technology to screen SUMO targets in a high-throughput fashion.

Characterizing the Selectivity and Specificity of SUMO Target Sites—For the simultaneous characterization of different SUMO target sites, we selected a panel of 11 validated SUMO substrates (Table 2). These targets included proteins involved in various different cellular processes, including signal transduction, gene transcription and DNA repair. For each wild-type peptide, the corresponding negative control peptide was also synthesized.

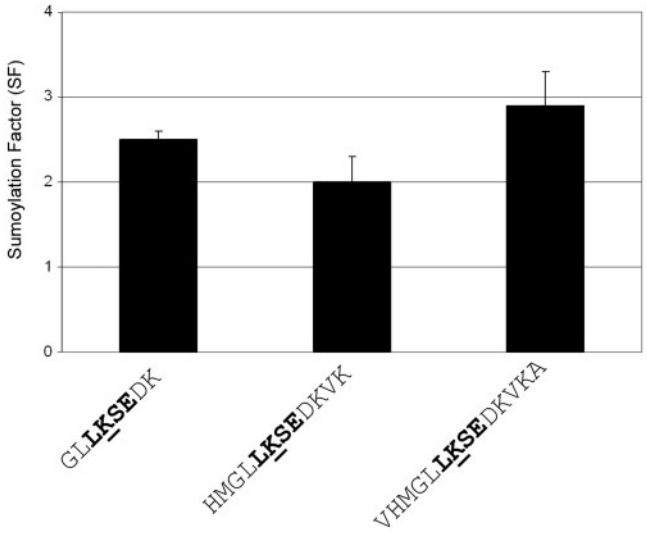

Fig. 3. **SUMO assay on Pepscan cards.** SUMO assay with peptides directly synthesized on a solid support (Pepscan cards) performed in optimized conditions: 0.15 μM E1, 0.20 μM E2 and 0.4 μM SUMO-1 in the presence of 1% BSA and 3% Tween-20 for 30 min at 37°C. Before detection, modified peptide cards were incubated in a sonication bath for 2 min (standard conditions). The amino acid residues of the RanGAP1 SUMO motif are indicated in bold letters and the acceptor Lys is underlined. Along the *y*-axis sumoylation factor (SF) values are given.

A comparison of peptides of 9- and 15-amino-acid residues in length revealed differences in the sumoylation patterns of the chosen SUMO substrates (Fig. 4A). For a few substrates, the longer version of the peptide gave a higher SF, whereas the reverse was true for other targets. Several peptides of 15amino-acid length showed higher sumoylation (*i.e.* HDAC-4, p53, SMAD4 and IκBα) than the shorter versions. In all cases, except for Sp100, the higher SF may be caused by negative charges in the longer C-terminus (Table 2) (*19*). Under these conditions, the target sites of HDAC4 and PML were more efficient substrates than the well-characterized RanGAP1 site. Moreover, our results reveal that PML is in general a better SUMO substrate than p53 and IκBα are consistent with other reports (*16*). As the observed sumoylation pattern for substrates between 9 and 15 amino acids in length were heterogeneous, we chose to use the longer version in subsequent studies, based on the assumption that the additional length may contribute additional information. In order to assure that the observed signals are based on covalent SUMO modifications, we tested the peptide substrates in the presence and absence of ATP. In the absence of ATP, we only detected a low background signal for the substrates (Supplementary Figure), ruling out the possibility of a non-covalent binding between SUMO and the peptides.

SUMO sites are often found at protein termini and extended loops (*36*, *37*); therefore, we compared cyclic peptides, potentially mimicking structural elements in the peptide, with linear peptides. Assuming that a conformational change occurred for all the cyclic peptides investigated, only the target site of PML

Table 2. **Peptides corresponding to SUMO target sites.**

	Protein	Position	Peptide sequence (wt)	Peptide sequence (ctrl)
1	HDAC-4	559	GVQVKQEPI	GVQVRQEPI
2	PML	490	RKVIKMESE	RKVIRMESE
3	Elk-1	249	PPEVKVEGP	PPEVRVEGP
4	p53	386	KLMFKTEGP	KLMFRTEGP
5	RanGAP1	526	MGLLKSEDK	MGLLRSEDK
6	SMAD4	159	SMMVKDEYV	SMMVRDEYV
7	IkBα	21	RDGLKKERL	RDGLRKERL
8	Sp100	297	LVDIKKEKP	LVDIRKEKP
9	c-Jun	229	LQALKEEPQ	LQALREEPQ
10	PIASy	35	KSGLKHELV	KSGLRHELV
11	NEMO	309	ADIYKADFQ	ADIYRADFQ
12	HDAC-4	559	AQAGVQVKQEPIESD	AQAGVQVRQEPIESD
13	PML	490	QCPRKVIKMESEEGK	QCPRKVIRMESEEGK
14	Elk-1	249	RALPPEVKVEGPKEE	RALPPEVRVEGPKEE
15	p53	386	RHKKLMFKTEGPDSD	RHKKLMFRTEGPDSD
16	RanGAP1	526	LIHMGLLKSEDKIKA	LIHMGLLRSEDKIKA
17	SMAD4	159	APSSMMVKDEYVHDF	APSSMMVRDEYVHDF
18	IkBα	21	EGPRDGLKKERLLDD	EGPRDGLRKERLLDD
19	Sp100	297	SVRLVDIKKEKPFSN	SVRLVDIRKEKPFSN
20	c-Jun	229	HPRLQALKEEPQTVP	HPRLQALREEPQTVP
21	PIASy	35	GRSKSGLKHELVTRA	GRSKSGLRHELVTRA
22	NEMO	309	KAQADIYKADFQAER	KAQADIYRADFQAER

The position indicates the amino acid position of the acceptor lysine in the full-length protein. The acceptor Lys and corresponding Arg in the control peptides are underlined. Wt (Wild-type), ctrl (control).

displayed a difference in SUMO modification pattern as a result of this change. The cyclic PML peptide had a higher SF than the corresponding linear peptides (Fig. 4B), suggesting that the conformational constraint or folding can influence the sumoylation efficiency of this target.

SUMO E3 ligases enhance the sumoylation of substrates *in vitro* and *in vivo*, but little is known about their mode of action. We investigated whether a SUMO E3 ligase could enhance peptide modification, by adding recombinant PIAS1 to the reaction mixture. The sumoylation of SMAD4, a known PIAS1 substrate (*38*) was specifically enhanced, by a factor of about 3 (Fig. 4C), indicating that the SUMO E3 ligase can fulfil its functions without a full-length target substrate. The sumoylation of p53, another PIAS1 target substrate (*39–41*) was only slightly enhanced in the presence of the ligase. Elk-1 has been identified as a substrate of PIASxα rather as a specific substrate of PIAS1 (*42*), but a slightly higher SF was obtained in the presence of PIAS1, suggesting that Elk-1 may be a substrate of PIAS1. None of the other substrates showed any significant change in the presence of PIAS1.

Cell extract may serve as another source of E3 ligases or proteins capable of enhancing or inhibiting the SUMO reaction. We therefore prepared cell lysate from HeLa cells and added it to the reaction mixture. For the SUMO sites of HDAC4, PML, PIASy and RanGAP1, the addition of cell lysate strongly decreased SF (Fig. 4D), suggesting that SUMO proteases present in the lysate inhibit the modification of these targets. Conversely, the sumoylation of p53 doubled, indicating enhancement mediated by an E3 ligase or another auxiliary protein. The other substrates displayed no significant difference in sumoylation in the presence and absence of cell lysate. The heterogeneous sumoylation pattern of the targets is consistent with a selective protein requirement for each individual substrate.

Deciphering the SUMO Consensus Site Motif—We investigated the composition of the consensus motif ψKxE/D in detail, by synthesizing a peptide library containing all possible amino acid combinations, with eight different hydrophobic amino acids in the first position, a Lys in the second position, all 20 amino acids in the third position and either Asp or Glu in the fourth position. All combinations, together with the control peptides, resulted in a library of 640 peptides. Figure 5A shows the Pepscan cards after detection of the SUMO-modified peptides. After calculating the SF for all peptides, the motifs were classified by their amino acid composition (Fig. 5B). The overall signal distribution highlights the SUMO efficiency of each individual target peptide and suggest preferences for each position in the consensus motif (see also Supplementary Data). Hence, a SF cut-off value of two was used to rank, for 107 peptides, the frequency of each amino acid at each position (Fig. 5C). The amino acid distribution for the first position showed a preference for Trp and Phe (upper left panel). A high frequency of hydrophobic residues, in particular aromatic amino acids, was also observed in the third position. Among the ranked peptides, basic and acidic amino acids were highly infrequent or did not occur at all in this position, indicating the avoidance of charged amino acids at this position (upper right panel). At the same position, the small amino acids Gly and Pro were also underrepresented. Glu was about twice as frequent as Asp in the fourth position in the consensus sequence (lower panel).

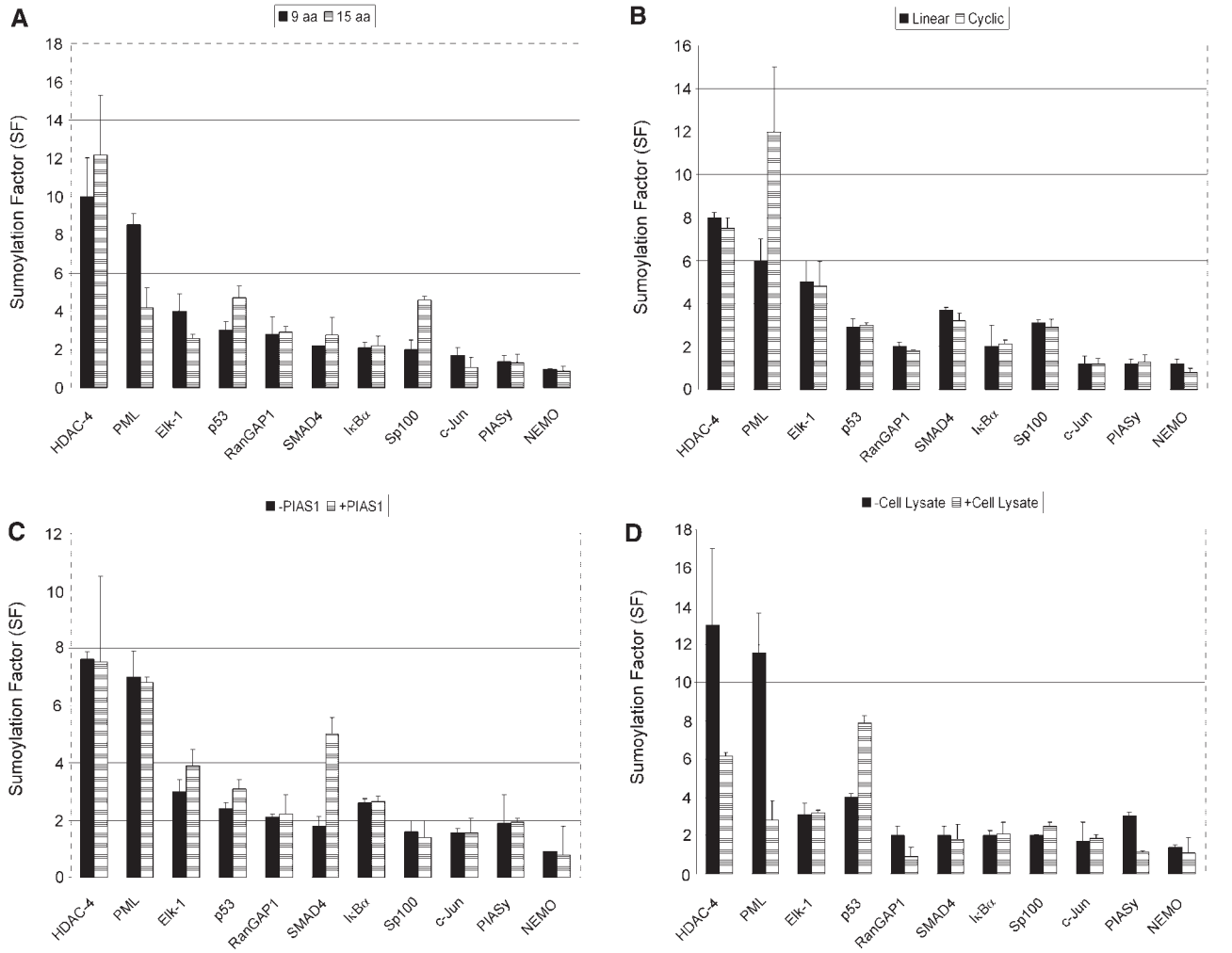

Fig. 4. **Characterization of SUMO target sites.** The SUMO assay was performed in standard conditions. (A) Comparison of 9- and 15-amino acid residue peptides. (B) Comparison of linear and cyclic 15-amino acid residue peptides. (C) SUMO reaction in the presence and absence of recombinant PIAS1 (500 nM) using 15-amino-acid residue peptides. (D) SUMO reaction in the presence and absence of HeLa total cell extracts (2.5 μg/μl) using 15-amino-acid residue peptides.

DISCUSSION

Sumoylation is a post-translational modification with a profound impact on almost all cellular processes (*2*, *3*, *43*, *44*). SUMO modification is a reversible process, facilitating the participation of proteins regulated in this manner in multiple rounds of functional circuits (*45*). Most SUMO substrates have the core consensus motif ψKxE/D in common. SUMO target sites in a given protein were identified by searching a protein database for the ψKxE/D motif. They were then confirmed or excluded based on *in vitro* sumoylation experiments. Mass spectrometry of *in vitro* sumoylated proteins can be used as an alternative (*25*, *46*). The detection of modified substrates and target sites is time-consuming, involving polyacrylamide gel electrophoresis followed by Coomassie blue staining for direct visualization or indirect visualization by immunoblotting, making it difficult to evaluate larger substrate sets. We report here the establishment of an *in vitro* SUMO assay with peptide arrays on a solid support and demonstrate that this assay format is useful for the characterization of a large number of different SUMO target sites simultaneously.

The length and amino acid composition of the investigated peptides generally affected the sumoylation reaction (Figs 3 and 4). A direct comparison of 11 different SUMO targets demonstrated that the substrates were modified with different efficiencies, suggesting a high degree of selectivity. Although the assay was optimized using the RanGAP1 protein, the target sites of HDAC4, PML and Elk-1 were particularly efficient SUMO substrates, as indicated by the higher SF value obtained for these sites. In addition to the classical consensus motif ψKxE/D, these three substrates have two acidic amino acid residues C-terminal from the core motif, consistent with recent reports of the existence of an extended negatively charged motif, NSDM, increasing the efficiency of the SUMO reaction (*19*).

Twenty-three percent of all validated sumoylation sites do not follow the consensus motif (*24*). Hence, we tested

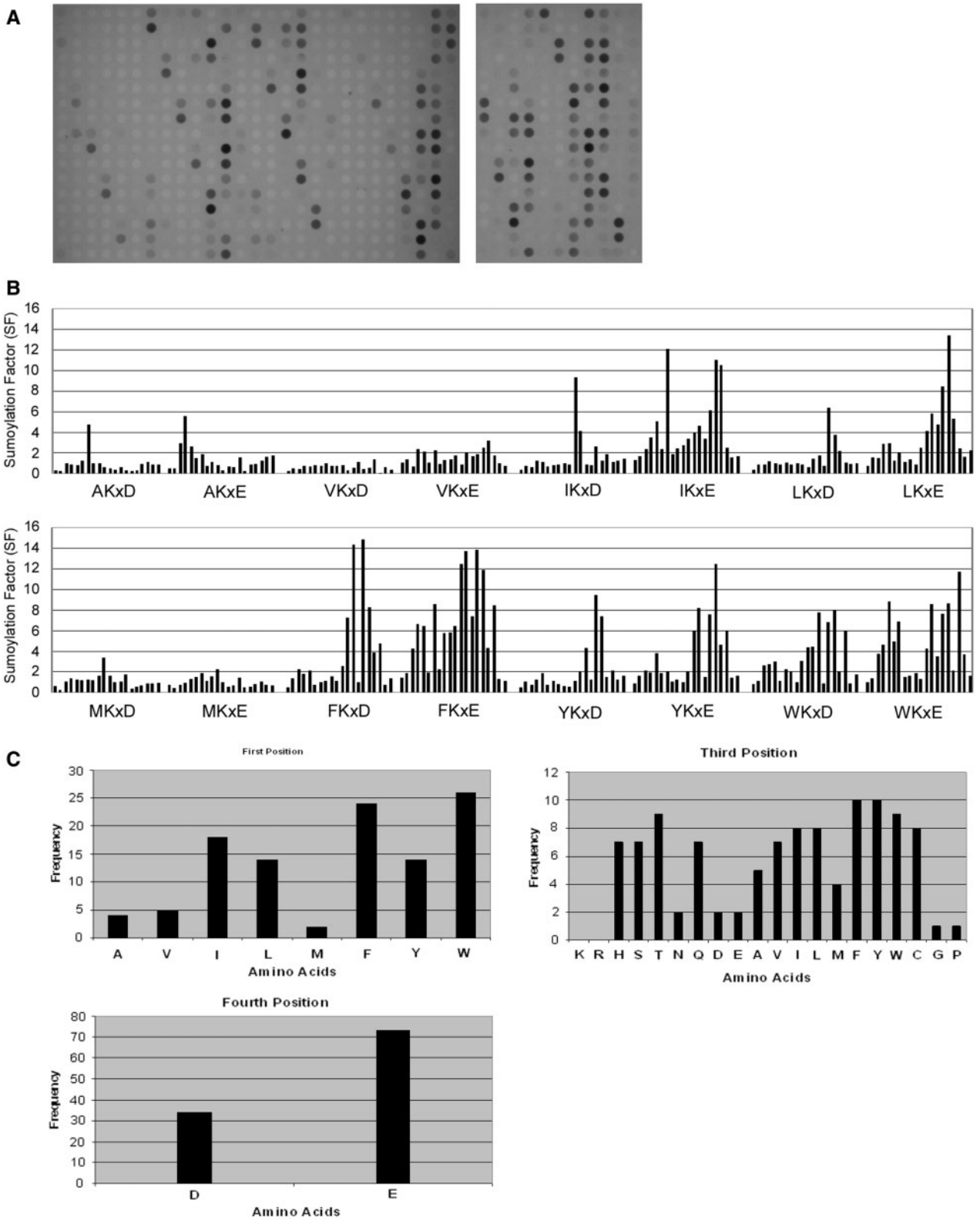

Fig. 5. **Deciphering the SUMO consensus site motif.** The SUMO assay was performed in standard conditions. (A) Picture of Pepscan cards after detection. Each spot represents an individual SUMO reaction in one well. (B) Overall signal distribution of all peptides. Each column corresponds to a specific target peptide with lysine in position 2 of the consensus motif. The *x*-axis represents the different consensus motifs on the card and the *y*-axis depicts the SF values for each peptide. *x* at position 3 in the motif indicates the amino acids in the following order (same order as in C, lower panel): Lys, Arg, His, Ser, Thr, Asn, Gln, Asp, Glu, Ala, Val Ile, Leu, Met, Phe, Tyr, Trp, Cys, Gly, Pro. (C) Assay analysis showing the frequency of each amino acid residue in positions 1 (upper left panel), 3 (upper right panel) and 4 (lower panel).

also non-consensus SUMO target sequences in this system: QAEAKCPKL (amino acids 61–69 of PML) and ISCAKDGVK (amino acids 160–168 of PCNA). However, these sites were only very weak SUMO substrates (data not shown), suggesting that for unstructured peptides the ψKxE/D motif is indeed the optimal target site motif.

Sumoylation, unlike ubiquitination, does not require an E3 ligase enzyme for completion of the transfer of SUMO to the substrate protein. Ubc9 directly binds SUMO substrates, as shown by the occurrence of sumoylation in the absence of E3 in a totally reconstituted *in vitro* system (*14*, *16*, *17*). However, the addition of an E3 enzyme specific to the substrate generally increases the rate and intensity of sumoylation. Interestingly, we found that PIAS1 specifically enhanced the sumoylation of SMAD4, a PIAS1-dependent substrate (Fig. 4C) (*38*). Moreover, the addition of HeLa cell lysate in the reaction mixture increased slightly p53 peptide sumoylation (Fig. 4D), suggesting the action of a SUMO E3 ligase *e.g.* PIAS proteins or TOPORS (*34*, *47–50*). But how can a SUMO E3 ligase increase the sumoylation of a 15-amino acid peptide? Several studies suggest that the E3 ligase does not bind the target, but only the Ubc9∼SUMO thioester intermediate to properly align Ubc9 and SUMO for the ligation reaction ('catalysis by proximity') (*36*, *51*). Although a contact between E3 and target can not be excluded, the very small SMAD4 peptide gives even stronger support to the model that E3 does not bind the target, but merely stimulate SUMO attachment to Lys residues that are selected by Ubc9 (*52*). Alternatively, an allosteric effect changes the specificity of Ubc9 to the target, including SMAD4, p53 and other substrates. Experiments are in progress to investigate in more detail the effect of SUMO E3 ligases on peptide sumoylation.

A recent search of the SWISSPROT database for human proteins with the ψKxE core motif identified 8,398 potential SUMO sites (*53*). Therefore, further details of the composition of this motif would be useful for target site prediction and would provide additional information about the interaction of the target site with Ubc9. The occurrence of residues at particular position of the tetrapeptide in the sumoylation motif ψKxE/D agrees with previous studies (Fig. 5C) (*16*, *17*). Our data show a relatively strong preference for the aromatic residues Phe and Trp for the first position. Although it is known that large hydrophobic residues are preferred in this position, the frequency in this study is higher than in the SUMO sites described until now. This may be caused by the fact that the substrate that was used was so minimal that the aromatic residues could provide more binding energy to Ubc9. Moreover, aromatic residues are mostly located in the interior of proteins and are underrepresented on the surface of proteins. Therefore, Phe and Trp may work very well in synthetic SUMO substrates but the occurrence in nature on an exposed motif is low. In addition, our finding that Ala in position 1 decreased the SUMO modification is in line with data from the same groups (*16*, *17*). The preference for position 3 is rather striking. According to the X-ray structure the side chain of residue three in the motif is pointed away from Ubc9 and the enzyme does not contact the side chain (*54*). However, in case of synthetic peptides, the side chain at the third position may have a structural influence on the folding of the peptide (for Gly and Pro) or form a hydrophobic scaffold that supports the structure (for Ile, Leu, Phe, Tyr and Trp). Surprisingly, among the 107 ranked peptides the basic amino acid residues Arg and Lys are completely avoided in position 3 in the consensus motif, although known validated SUMO target sites like IκBα and SP100 contain an Arg or Lys at this position. However, basic amino acids at position 3 were found in peptides with a SF <2, indicating that Arg and Lys might contribute to a non-covalent SUMO binding.

The here-described assay format could be used to test the other members of the SUMO family and to compare their target specificities. The assay could also provide a snapshot of SUMO targets in cancer cells. The role of modified SUMO substrates in cancer development could be investigated by testing cell lysates from cancer and non-cancer cells in this assay. The sumoylation pattern obtained would indicate which substrates were the most likely to be involved in the studied cancers. The high-throughput assay format is particularly suitable for the screening of chemical compounds potentially affecting the modification of therapeutic SUMO targets. Furthermore, the combination of search engines such as SUMOplot (http://www.abgent.com/doc/sumoplot) or SUMOsp (http://bioinformatics.lcd-ustc.org/sumosp/) with this assay should help to accelerate the identification of SUMO target sites; however, follow-up experiments in the full-length protein are still required for validation.

In summary, the *in vitro* sumoylation assay with peptide arrays on a solid support provides a robust and sensitive method for the simultaneous analysis of a large number of SUMO target sites. It can be used to determine the selectivity of SUMO target sites and for investigating the specificity of E3 ligases.

Supplementary data are available at *JB* online.

We thank Alfred C.O. Vertegaal for helpful comments on the manuscript and Joshua Buijnink for technical assistance. This work was supported by grants from EEC (6th program, RUBICON).

REFERENCES

1. Hay, R.T. (2005) SUMO: a history of modification. *Mol. Cell* **18**, 1–12
2. Johnson, E.S. (2004) Protein modification by SUMO. *Annu. Rev. Biochem.* **73**, 355–382
3. Seeler, J.S. and Dejean, A. (2003) Nuclear and unclear functions of SUMO. *Nat. Rev. Mol. Cell. Biol.* **4**, 690–699
4. Li, Y., Wang, H., Wang, S., Quon, D., Liu, Y.W., and Cordell, B. (2003) Positive and negative regulation of APP amyloidogenesis by sumoylation. *Proc. Natl. Acad. Sci. USA* **100**, 259–264
5. Steffan, J.S., Agrawal, N., Pallos, J., Rockabrand, E., Trotman, L.C., Slepko, N., Illes, K., Lukacsovich, T., Zhu, Y.Z., Cattaneo, E., Pandolfi, P.P., Thompson, L. M., and Marsh, J.L. (2004) SUMO modification of Huntingtin and Huntington's disease pathology. *Science* **304**, 100–104
6. Guo, D., Han, J., Adam, B.L., Colburn, N.H., Wang, M.H., Dong, Z., Eizirik, D.L., She, J.X., and Wang, C.Y. (2005)

Proteomic analysis of SUMO4 substrates in HEK293 cells under serum starvation-induced stress. *Biochem. Biophys. Res. Commun.* **337**, 1308–1318
7. Li, T., Evdokimov, E., Shen, R.F., Chao, C.C., Tekle, E., Wang, T., Stadtman, E. R., Yang, D.C., and Chock, P.B. (2004) Sumoylation of heterogeneous nuclear ribonucleoproteins, zinc finger proteins and nuclear pore complex proteins: a proteomic analysis. *Proc. Natl. Acad. Sci. USA* **101**, 8551–8556
8. Vertegaal, A.C., Andersen, J.S., Ogg, S.C., Hay, R.T., Mann, M., and Lamond, A.I. (2006) Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. *Mol. Cell Proteomics* **5**, 2298–2310
9. Vertegaal, A.C., Ogg, S.C., Jaffray, E., Rodriguez, M.S., Hay, R.T., Andersen, J.S., Mann, M., and Lamond, A.I. (2004) A proteomic study of SUMO-2 target proteins. *J. Biol. Chem.* **279**, 33791–33798
10. Zhao, Y., Kwon, S.W., Anselmo, A., Kaur, K., and White, M.A. (2004) Broad spectrum identification of cellular small ubiquitin-related modifier (SUMO) substrate proteins. *J. Biol. Chem.* **279**, 20999–21002
11. Kagey, M.H., Melhuish, T.A., and Wotton, D. (2003) The polycomb protein Pc2 is a SUMO E3. *Cell* **113**, 127–137
12. Pichler, A., Gast, A., Seeler, J.S., Dejean, A., and Melchior, F. (2002) The nucleoporin RanBP2 has SUMO1 E3 ligase activity. *Cell* **108**, 109–120
13. Sachdev, S., Bruhn, L., Sieber, H., Pichler, A., Melchior, F., and Grosschedl, R. (2001) PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies. *Genes Dev.* **15**, 3088–3103
14. Bernier-Villamor, V., Sampson, D.A., Matunis, M.J., and Lima, C.D. (2002) Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* **108**, 345–356
15. Lin, D., Tatham, M.H., Yu, B., Kim, S., Hay, R.T., and Chen, Y. (2002) Identification of a substrate recognition site on Ubc9. *J. Biol. Chem.* **277**, 21740–21748
16. Rodriguez, M.S., Dargemont, C., and Hay, R.T. (2001) SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *J. Biol. Chem.* **276**, 12654–12659
17. Sampson, D.A., Wang, M., and Matunis, M.J. (2001) The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. *J. Biol. Chem.* **276**, 21664–21669
18. Subramanian, L., Benson, M.D., and Iniguez-Lluhi, J.A. (2003) A synergy control motif within the attenuator domain of CCAAT/enhancer-binding protein alpha inhibits transcriptional synergy through its PIASy-enhanced modification by SUMO-1 or SUMO-3. *J. Biol. Chem.* **278**, 9134–9141
19. Yang, S.H., Galanis, A., Witty, J., and Sharrocks, A.D. (2006) An extended consensus motif enhances the specificity of substrate modification by SUMO. *EMBO J.* **25**, 5083–5093
20. Yang, S.H., Jaffray, E., Hay, R.T., and Sharrocks, A.D. (2003) Dynamic interplay of the SUMO and ERK pathways in regulating Elk-1 transcriptional activity. *Mol. Cell* **12**, 63–74
21. Hietakangas, V., Anckar, J., Blomster, H.A., Fujimoto, M., Palvimo, J.J., Nakai, A., and Sistonen, L. (2006) PDSM, a motif for phosphorylation-dependent SUMO modification. *Proc. Natl. Acad. Sci. USA* **103**, 45–50
22. Pichler, A., Knipscheer, P., Oberhofer, E., van Dijk, W.J., Korner, R., Olsen, J.V., Jentsch, S., Melchior, F., and Sixma, T.K. (2005) SUMO modification of the ubiquitin-conjugating enzyme E2-25K. *Nat. Struct. Mol. Biol.* **12**, 264–269
23. Liu, B., Li, S., Wang, Y., Lu, L., Li, Y., and Cai, Y. (2007) Predicting the protein SUMO modification sites based on Properties Sequential Forward Selection (PSFS). *Biochem. Biophys. Res. Commun.* **358**, 136–139
24. Xue, Y., Zhou, F., Fu, C., Xu, Y., and Yao, X. (2006) SUMOsp: a web server for sumoylation site prediction. *Nucleic Acids Res.* **34**, W254–W257
25. Pedrioli, P.G., Raught, B., Zhang, X.D., Rogers, R., Aitchison, J., Matunis, M., and Aebersold, R. (2006) Automated identification of SUMOylation sites using mass spectrometry and SUMmOn pattern recognition software. *Nat. Methods* **3**, 533–539
26. Oh, Y.H., Hong, M.Y., Jin, Z., Lee, T., Han, M.K., Park, S., and Kim, H.S. (2007) Chip-based analysis of SUMO (small ubiquitin-like modifier) conjugation to a target protein. *Biosens. Bioelectron.* **22**, 1260–1267
27. Martin, S.F., Hattersley, N., Samuel, I.D., Hay, R.T., and Tatham, M.H. (2007) A fluorescence-resonance-energy-transfer-based protease activity assay and its use to monitor paralog-specific small ubiquitin-like modifier processing. *Anal. Biochem.* **363**, 83–90
28. Bossis, G., Chmielarska, K., Gartner, U., Pichler, A., Stieger, E., and Melchior, F. (2005) A fluorescence resonance energy transfer-based assay to study SUMO modification in solution. *Methods Enzymol.* **398**, 20–32
29. Timmerman, P., Beld, J., Puijk, W.C., and Meloen, R.H. (2005) Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces. *Chembiochem* **6**, 821–824
30. Geysen, H.M., Meloen, R.H., and Barteling, S.J. (1984) Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc. Natl. Acad. Sci. USA* **81**, 3998–4002
31. Slootstra, J.W., Puijk, W.C., Ligtvoet, G.J., Langeveld, J.P., and Meloen, R.H. (1996) Structural aspects of antibody-antigen interaction revealed through small random peptide libraries. *Mol. Divers.* **1**, 87–96
32. Mahajan, R., Delphin, C., Guan, T., Gerace, L., and Melchior, F. (1997) A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* **88**, 97–107
33. Okuma, T., Honda, R., Ichikawa, G., Tsumagari, N., and Yasuda, H. (1999) *Biochem. Biophys. Res. Commun.* **254**, 693–698
34. Bischof, O., Schwamborn, K., Martin, N., Werner, A., Sustmann, C., Grosschedl, R., and Dejean, A. (2006) *Mol. Cell* **22**, 783–794
35. Yunus, A.A. and Lima, C.D. (2005) Purification and activity assays for Ubc9, the ubiquitin-conjugating enzyme for the small ubiquitin-like modifier SUMO. *Methods Enzymol.* **398**, 74–87
36. Dye, B.T. and Schulman, B.A. (2007) Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 131–150
37. Macauley, M.S., Errington, W.J., Scharpf, M., Mackereth, C.D., Blaszczak, A. G., Graves, B.J., and McIntosh, L.P. (2006) Beads-on-a-string, characterization of ETS-1 sumoylated within its flexible N-terminal sequence. *J. Biol. Chem.* **281**, 4164–4172
38. Liang, M., Melchior, F., Feng, X.H., and Lin, X. (2004) Regulation of Smad4 sumoylation and transforming growth factor-beta signaling by protein inhibitor of activated STAT1. *J Biol. Chem.* **279**, 22857–22865
39. Kahyo, T., Nishida, T., and Yasuda, H. (2001) Involvement of PIAS1 in the sumoylation of tumor suppressor p53. *Mol. Cell* **8**, 713–718
40. Megidish, T., Xu, J.H., and Xu, C. W. (2002) Activation of p53 by protein inhibitor of activated Stat1 (PIAS1). *J. Biol. Chem.* **277**, 8255–8259

41. Schmidt, D. and Muller, S. (2002) Members of the PIAS family act as SUMO ligases for c-Jun and p53 and repress p53 activity. *Proc. Natl. Acad. Sci. USA* **99**, 2872–2877
42. Yang, S.H. and Sharrocks, A.D. (2005) PIASx acts as an Elk-1 coactivator by facilitating derepression. *EMBO J.* **24**, 2161–2171
43. Melchior, F. (2000) SUMO–nonclassical ubiquitin. *Annu. Rev. Cell. Dev. Biol.* **16**, 591–626
44. Muller, S., Hoege, C., Pyrowolakis, G., and Jentsch, S. (2001) SUMO, ubiquitin's mysterious cousin. *Nat. Rev. Mol. Cell Biol.* **2**, 202–210
45. Kerscher, O., Felberbaum, R., and Hochstrasser, M. (2006) Modification of proteins by ubiquitin and ubiquitin-like protein. *Annu. Rev. Cell. Dev. Biol.* **22**, 159–180
46. Knuesel, M., Cheung, H.T., Hamady, M., Barthel, K.K., and Liu, X. (2005) A method of mapping protein sumoylation sites by mass spectrometry using a modified small ubiquitin-like modifier 1 (SUMO-1) and a computational program. *Mol. Cell Proteomics* **4**, 1626–1636
47. Gostissa, M., Hengstermann, A., Fogal, V., Sandy, P., Schwarz, S.E., Scheffner, M., and Del Sal, G. (1999) Activation of p53 by conjugation to the ubiquitin-like protein SUMO-1. *EMBO J.* **18**, 6462–6471
48. Muller, S., Berger, M., Lehembre, F., Seeler, J.S., Haupt, Y., and Dejean, A. (2000) c-Jun and p53 activity is modulated by SUMO-1 modification. *J. Biol. Chem.* **275**, 13321–13329
49. Pungaliya, P., Kulkarni, D., Park, H.J., Marshall, H., Zheng, H., Lackland, H., Saleem, A., and Rubin, E.H. (2007) TOPORS functions as a SUMO-1 E3 ligase for chromatin-modifying proteins. *J. Proteome Res.*
50. Weger, S., Hammer, E., and Heilbronn, R. (2005) Topors acts as a SUMO-1 E3 ligase for p53 in vitro and in vivo. *FEBS Lett.* **579**, 5007–5012
51. Reverter, D. and Lima, C.D. (2005) Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature* **435**, 687–692
52. Reindle, A., Belichenko, I., Bylebyl, G.R., Chen, X.L., Gandhi, N., and Johnson, E.S. (2006) Multiple domains in Siz SUMO ligases contribute to substrate selectivity. *J. Cell Sci.* **119**, 4749–4757
53. Yang, X.J. and Gregoire, S. (2006) A recurrent phosphosumoyl switch in transcriptional repression and beyond. *Mol. Cell* **23**, 779–786
54. Yunus, A.A. and Lima, C.D. (2006) Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway. *Nat. Struct. Mol. Biol.* **13**, 491–499