

Identification of Sumoylation Inhibitors Targeting a Predicted Pocket in Ubc9

Ashutosh Kumar,^{†,‡} Akihiro Ito,^{‡,§,⊥} Mikako Hirohama,^{‡,||} Minoru Yoshida,^{*,‡,§,||} and Kam Y. J. Zhang^{*,†}

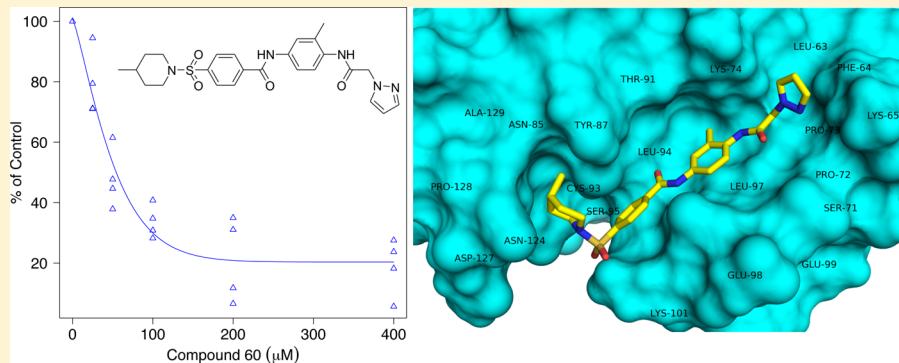
[†]Structural Bioinformatics Team, Division of Structural and Synthetic Biology, Center for Life Science Technologies, RIKEN, 1-7-22 Suehiro, Yokohama, Kanagawa 230-0045, Japan

[‡]Chemical Genetics Laboratory, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

[§]Chemical Genomics Research Group, RIKEN Center for Sustainable Resource Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

^{||}CREST Research Project, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

Supporting Information



ABSTRACT: Sumoylation is a post-translational modification that plays an important role in a wide range of cellular processes. Among the proteins involved in the sumoylation pathway, Ubc9 is the sole E2-conjugating enzyme required for sumoylation and plays a central role by interacting with almost all of the partners required for sumoylation. Ubc9 has been implicated in a variety of human malignancies. In order to exploit the therapeutic potential of Ubc9, we have identified the potential site to target for rational drug design using molecular modeling approaches. The structural information derived was then used to prioritize hits from a small-molecule library for biological assay using a virtual screening protocol that involves shape matching with a known inhibitor inhibitors and docking of a small-molecule library utilizing computational approaches that incorporate both ligand and protein flexibility. Nineteen compounds were acquired from different chemical vendors and were tested for Ubc9 inhibitory activity. Five compounds showed inhibitory activity against Ubc9, out of which one compound was selected for further optimization. A similarity search was then carried out to retrieve commercially available derivatives, which were further acquired and assayed, resulting in two compounds with acceptable potency. These two compounds can be used as starting points for the development of more potent inhibitors of Ubc9 targeting the predicted site.

INTRODUCTION

Sumoylation is a post-translational modification that is critical to cellular processes such as DNA replication and repair, chromosome packing and dynamics, genome integrity, nuclear transport, signal transduction, and cell proliferation.^{1–7} In sumoylation, a small ubiquitin-like modifier (SUMO) protein is covalently attached to the ε-amino group of lysine residues in specific target proteins.^{5,8–10} This conjugation requires sequential action of a set of three enzymes: an activating enzyme E1 (SUMO E1), a conjugating enzyme E2 (SUMO E2 or Ubc9), and a ligase E3 (SUMO E3).^{5,8–10} Among these proteins, Ubc9 is required for the conjugation of every SUMO paralogue to substrate proteins.^{11–13} Ubc9 is the only known SUMO-conjugating enzyme required for sumoylation, and its deletion abolishes SUMO conjugation.⁹ Over the past few years, several

studies have implicated Ubc9 in a variety of human diseases, including cancer, neurodegenerative diseases, and heart diseases.^{14–22} It has been observed that Ubc9 is overexpressed in several malignancies, such as ovarian carcinoma,^{15,22} melanoma,^{23,24} and lung adenocarcinoma. The overexpression of Ubc9 can increase cancer cell growth.^{15,25} Furthermore, Ubc9 was one of the down-regulated proteins in a proteomic study where p53 up-regulated modulator of apoptosis (PUMA), a BH3-only pro-apoptotic protein that induces BAX-dependent apoptosis, was overexpressed.²⁶ Apart from p53, Ubc9 also regulates the activity of other important tumor suppressor proteins, including retinoblastoma protein (pRB), p63, p73, and murine double

Received: July 7, 2014

Published: September 5, 2014

minute 2 (MDM2).²⁷ These studies suggest that Ubc9 may be fundamental for tumorigenesis and tumor progression by preventing the activation of apoptotic pathways. Ubc9 is also involved in certain heart diseases such as cardiomyopathies, muscular diseases such as muscular dystrophy, and neurodegenerative diseases such as Huntington's, Alzheimer's, and Parkinson's diseases.¹⁷ In a yeast two-hybrid screening study, an interaction between lamin A protein and Ubc9 was detected. Lamin A protein is an important regulator of nuclear structure and function, and mutations in lamin A protein cause various diseases such as cardiomyopathy and muscular dystrophy.²⁸ Also, Ubc9 is known to interact with a number of neurodegenerative-disease-related proteins such as huntingtin (Huntington's disease), ataxin-1 (spinocerebellar ataxia type 1), tau, α -synuclein, DJ-1 (Parkinson's disease), and APP (Alzheimer's disease).¹⁷

The above studies suggest that Ubc9 is an attractive target for drug development against a variety of human diseases. However, only a few natural products have been identified as Ubc9 inhibitors to date.^{29,30} These include flavonoid derivatives reported by Kim et al.²⁹ using an electrophoretic mobility shift assay. The reported compounds inhibited the transfer of the SUMO from the Ubc9 thioester conjugate to the substrate protein. The synthesis of the most potent compound was also reported in another study.³¹ Using high-throughput screening of a natural product library, Hirohama et al.³⁰ identified spectomycin B1 as another potent inhibitor of Ubc9 that directly binds to SUMO E2 and selectively blocks the formation of the Ubc9-SUMO thioester intermediate. Although the reported natural products are fairly potent, they are less amenable for chemical optimization. To fully exploit the therapeutic potential of Ubc9 as a drug target, it is indispensable to identify more chemical scaffolds that are suitable for extensive chemical optimization for lead identification and development.

The design of small-molecule inhibitors against Ubc9 has been hampered by the lack of information about the potential site to target by a structure-based design approach. Although many sites have been suggested for small-molecule inhibitor discovery,¹⁸ no small-molecule inhibitors targeting these sites have been reported. Motivated by the need for druggable sites that can be targeted for the inhibition of Ubc9 biological activity, we have identified a site near the catalytic Cys93 as a potential site to target for structure-based drug design. The identification of druglike ligands with moderate inhibitory activity against Ubc9 further validates the druggability of this site.

MATERIALS AND METHODS

Pocket Prediction. Fifteen available human Ubc9 crystal structures (Table S1 in the Supporting Information) were retrieved from the RCSB Protein Data Bank (PDB),³² and all of the protein structures were prepared for pocket prediction using the protein preparation utility in Maestro.³³ All of the crystal structures were superimposed on wild-type Ubc9 (PDB code 2GRN) prior to pocket prediction. The SiteMap program^{34–36} was employed to predict binding pockets on the surface of Ubc9 structures. Default SiteMap parameters were used for pocket prediction. The five top ranked binding pockets of each structure were identified. Sites with at least 15 points were selected. A standard grid of 1.0 Å and the OPLS2005 force field³⁷ were used. Graphics were prepared using PyMOL³⁸ and Gnuplot 4.6.³⁹

Shape Similarity Calculations. The ShaEP program⁴⁰ was employed to perform shape similarity calculations against the ZINC lead-like subset⁴¹ using the Ubc9 inhibitor spectomycin

B1³⁰ as a search query. ShaEP carries out rigid-body superimposition and calculates the similarity between two small molecules by maximizing the spatial overlap in volume and electrostatic potential. In this study, the similarity search was performed using the “onlyshape” option of ShaEP.

Preparation of Ligand and Receptor Structures for Virtual Screening. The ZINC lead-like subset was downloaded, and three-dimensional structures were regenerated using OMEGA version 2.4.^{42–44} The structures were generated using a modified version of the MMFF94 force field.⁴⁵ The maximum number of conformations per molecule (maxconfs) option was set to 1 in order to output the lowest-energy conformer only. An energy window for acceptable conformers (ewindow) of 10 kcal/mol and a root-mean-square deviation (RMSD) cutoff of 0.5 Å were used. The structures were then imported into Molecular Operating Environment (MOE)⁴⁶ to prepare the small-molecule library for virtual screening by assigning partial charges and generating the protonation and tautomeric states. Partial charges were assigned according to the AM1-BCC semiempirical method.⁴⁷ A Ubc9 crystal structure (PDB code 2GRN) was used for molecular docking. To prepare the receptor structure, hydrogens were added and bond orders were assigned using the protein preparation utility of MOE. The protonation states of charged residues were determined using Protonate3D.⁴⁸ For RosettaLigand docking,^{49–51} receptor side chains were repacked using the “ligand_rpkmin” protocol of the Rosetta suite⁵² to avoid pre-existing clashes according to the Rosetta energy function. RosettaLigand uses pregenerated ligand conformations, so a conformational ensemble of a maximum of 200 conformations for each ligand was generated using OMEGA^{42–44} for RosettaLigand docking. The same OMEGA parameters as mentioned above were used except that the maxconfs option was set to 200.

Molecular Docking. Docking-based virtual screening was performed using a two-step docking procedure. Initially, the Dock6.4 program^{53–56} was used to eliminate compounds with noncompatible geometries and energetics. Top-ranked compounds were then docked flexibly to Ubc9 using the RosettaLigand protocol in the Rosetta suite.^{49–51} For Dock6.4 docking, the binding site was defined using the Spghen, Sphere_selector, and Showsphere utilities of Dock6.4. Grids for docking were generated using a grid spacing of 0.5 Å, a distance cutoff of 9999 Å, and a distance-dependent dielectric constant value of 4 utilizing the Grid utility of Dock6.4. Docking was performed using the “flexible_ligand” option, and 20 poses per ligand were generated. The Dock6.4 scoring function “gridscore”⁵⁷ was used to rank-order the compounds, and the pose with the best gridscore was selected. For RosettaLigand docking, pregenerated ligand conformations were placed into the top-ranked predicted pocket (pocket A) of Ubc9, followed by the placement of amino acid side-chain rotamers around the ligand. Randomly sampled flexible ligand poses were then optimized using a Metropolis Monte Carlo simulated annealing algorithm. We generated approximately 5000 ligand poses, which were ranked by the Rosetta energy function. The best 5% of the poses were reranked on the basis of their InterfaceDelta scores (representing the difference between the energies of the protein in the ligand-bound and unbound states), and the pose with the lowest InterfaceDelta score was then chosen as the best-docked pose.

In Vitro Sumoylation Assay and SUMO1 Thioester Bond Formation Assay. The in vitro sumoylation reaction was performed as previously described.⁵⁸ Briefly, reaction mixtures

containing 0.1 μg of His and T7-tagged RanGAP1-C2, 0.3 μg of GST-Aos1/Uba2, 0.01 μg of His-tagged Ubc9, and 0.1 μg of GFP-fused SUMO-1 were incubated for 2 h at 30 °C. Samples were separated by 10% SDS-PAGE, and the fluorescence of GFP-fused sumoylated RanGAP1-C2 was analyzed with an Image-Quant LAS 4000 imager (GE Healthcare). The thioester bond formation reaction between SUMO-1 and E1 was performed as described previously.⁵⁸ To detect the thioester bond formation between SUMO-1 and E2, reaction mixtures containing 0.1 μg of biotinylated-SUMO-1, 0.3 μg of GST-Aos1/Uba2, and 0.5 μg of His-tagged Ubc9 were incubated for 30 min at 37 °C. Samples were separated by 12% SDS-PAGE, and the E2-biotinylated SUMO-1 intermediate was detected using avidin-conjugated horseradish peroxidase (Sigma).

Molecular Dynamics Simulations. Docking-predicted binding modes of compounds **25** and **60** were used as input structures for molecular dynamics (MD) simulations using the Desmond program.⁵⁹ The protein–ligand complex system for MD simulations was prepared using the “system builder” utility of Maestro.³³ Briefly, the protein–ligand complex was placed in an orthorhombic box ensuring 10 Å solvent buffers. The TI4P water model was used for MD simulations, and Na⁺ and Cl⁻ were added at a concentration of 0.15 M to neutralize the system. The parameters for simulations were assigned using the OPLS-AA/2005 force field.³⁷ Prior to the production run, the protein–ligand system was relaxed using a six-step relaxation protocol as implemented in Desmond. Finally, a 200 ns production simulation was performed using the *NPT* ensemble. The temperature was set at 300 K using a Nosé–Hoover thermostat. The SHAKE algorithm⁶⁰ was used to constrain all of the hydrogen atoms, and a time step of 5 fs was used. The long-range electrostatic interactions were handled using the particle-mesh Ewald (PME) method.⁶¹ A distance cutoff of 9.0 Å was used for short-range electrostatics and Lennard-Jones interactions. The trajectory was analyzed using Maestro,³³ and graphics were prepared using Gnuplot 4.6.³⁹

Hardware. The molecular docking calculations were performed on the 97.4 TFLOPS Intel Xeon 5570-based Massively Parallel PC Cluster of the RIKEN Integrated Cluster of Clusters (RICC). The results were analyzed on a Dell Precision T5400 workstation with a 2.0 GHz Intel Xeon CPU.

Compounds. Tested compounds were purchased from different chemical vendors via a local distributor. The vendors had verified that each compound had >95% purity by liquid chromatography–mass spectrometry and NMR experiments. Compounds were maintained as DMSO stock solutions.

■ RESULTS

Probing of Putative Binding Sites on the Ubc9 Surface.

Our investigation started with an exploration of all available human Ubc9 crystal structures (Table S1 in the Supporting Information) for potential small-molecule binding pockets. The Sitemap program^{34–36} was used to predict small-molecule binding pockets on the surface of 15 human Ubc9 crystal structures. Overall, seven binding pockets (sites A to G) were located on the surface of Ubc9 (Figure 1) with SiteScores ranging from 0.53 to 0.92 (Table S1 in the Supporting Information). Pockets A and B as well as pockets D and E were found to be very close and overlapping with each other, whereas C, F, and G were clearly distinct from the other predicted pockets. Pocket A, which was flanked by the residues Lys65, Ser71, Pro72, Pro73, Lys74, Cys75, Asn85, Tyr87, Thr91, Val92, Cys93, Leu94, Ser95, Leu97, Glu98, Gln99, Lys101, Asn124, Gln126, and Asp127

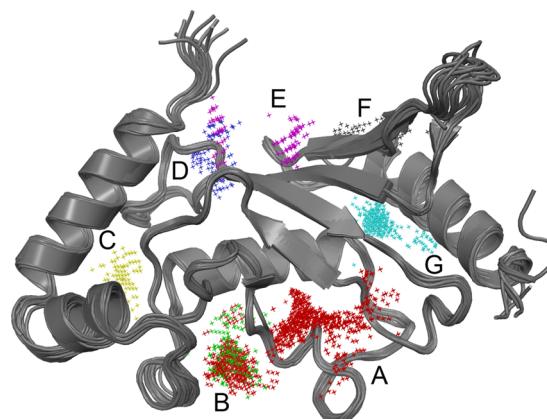


Figure 1. Sitemap-predicted pockets for 15 superimposed human Ubc9 crystal structures. Colored points represent the top SiteMap predictions for the 15 crystal structures: site A, red; site B, green; site C, yellow; site D, blue; site E, magenta; site F, dark gray; site G, cyan.

(Figure S1A in the Supporting Information), was predicted to be the top-ranked pocket on the surface of Ubc9 for the majority (11 out of 15) of the human Ubc9 crystal structures. This predicted pocket seems to be a good candidate for targeting with druglike small molecules. The majority of residues constituting this pocket are conserved among Ubc9 orthologs from various species (Figure S2A in the Supporting Information) and other analogous proteins such as ubiquitin-conjugating enzyme UbcH7 and Nedd8-conjugating enzyme Ubc12 (Figure S2B in the Supporting Information). Only this pocket has SiteScore values of >0.8 for nine out of 15 human Ubc9 crystal structures (Figure 2A) which is a recommended cutoff for druggability of small-molecule binding pockets.^{62,63} Although a large variation was found in the volume of pocket A among all of the crystal structures, it has an appropriate volume (~200 Å³) in a majority of the crystal structures (Figure 2B). Moreover, pocket A contains the important catalytic residue (Cys93) and constitutes the major contact region for intermediate covalent interaction between SUMO and Ubc9.⁶⁴ Additionally, this is very close to the binding site for sumoylation substrates such as RanGAP1.⁶⁴ Furthermore, pocket A has a higher fragment probe occupancy, as revealed from computational fragment mapping calculations using the FTMap program,⁶⁵ indicating a ligand-binding hotspot in this region (Figure S1B in the Supporting Information).

In Silico Virtual Screening To Assess the Druggability of the Predicted Pocket. To identify inhibitors targeting pocket A, we followed a hierarchical virtual screening protocol involving a combination of ligand shape similarity with two stages of flexible molecular docking calculations. Initially, the lead-like subset of the ZINC database⁴¹ was reduced by eliminating compounds distant from the Ubc9 inhibitor spectomycin B1³⁰ on the basis of their three-dimensional shapes. The ShaEP program⁴⁰ was used to carry out the shape calculations. A Tanimoto coefficient⁶⁶ cutoff of 0.6 resulted in 618 608 compounds, which were then docked to Ubc9 to prioritize the compounds on the basis of their ability to form favorable interactions in the binding pocket. Docking was performed following a two-step docking strategy starting with fast rigid receptor docking using Dock6.4.^{53–56} Rigid docking was followed by finer levels of sampling with more strict descriptions of the scoring functions implemented in the flexible docking program RosettaLigand.^{49–51} Our two-step docking strategy is a good trade-off between accuracy and computational power. A

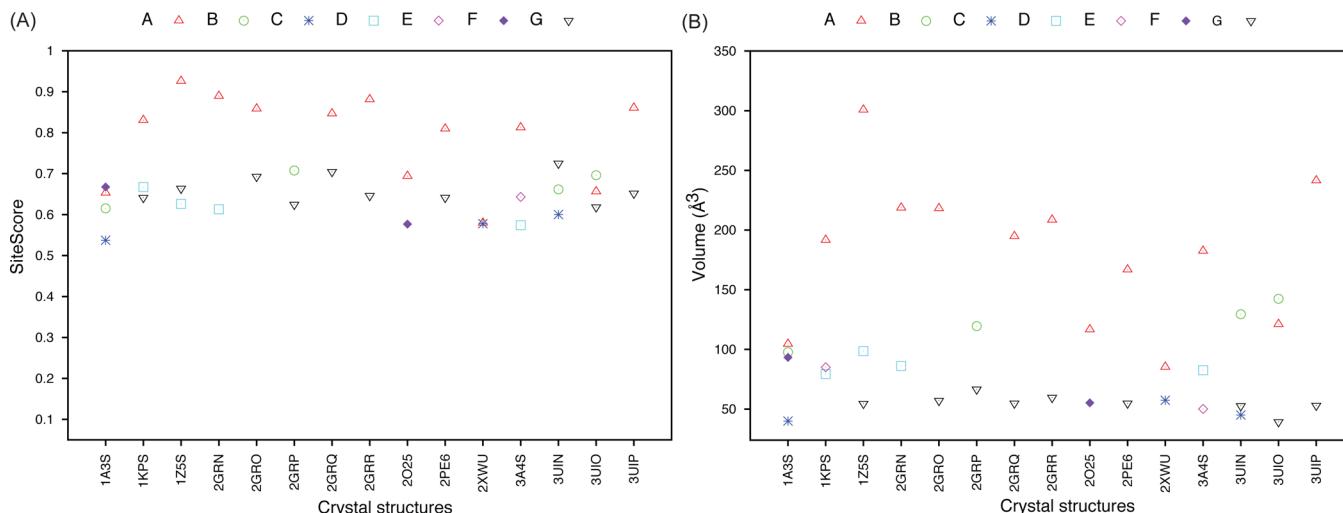


Figure 2. Properties of the seven SiteMap-predicted binding pockets (A to G, represented as points of different shapes and colors): (A) SiteScore; (B) volume.

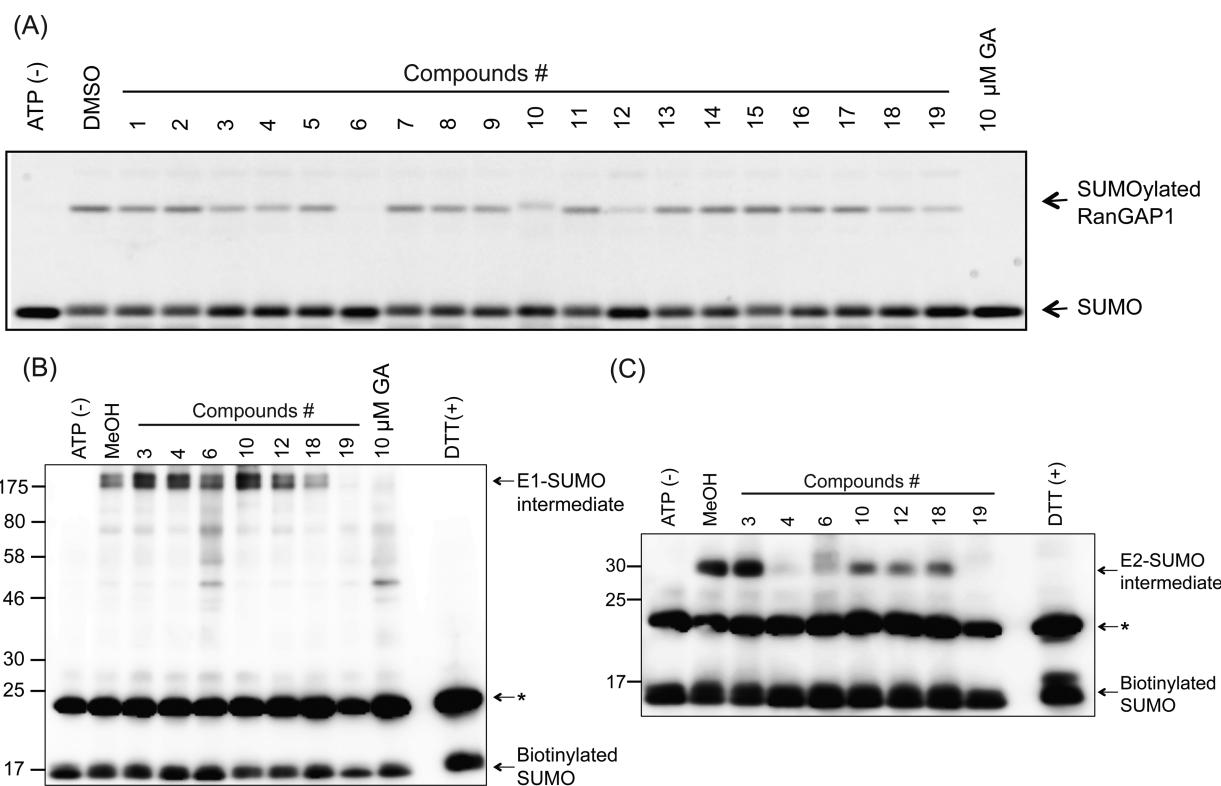


Figure 3. Ubc9 inhibitory activities of virtual screening hits. (A) In vitro protein sumoylation assay for 19 compounds. (B) SUMO E1-biotinylation intermediate formation assay for eight compounds showing inhibition of in vitro protein sumoylation. (C) SUMO E2-biotinylation intermediate formation assay for the same set of compounds as in (B). The asterisk (*) represents a non-specific band.

similar docking strategy has been utilized to identify inhibitors of SUMO activating enzyme 1.^{67,68} The utility of cascade or multistep hierarchical docking has also been reviewed previously.⁶⁹ In the first stage of this docking protocol, the 618 608 compounds obtained from the shape similarity calculations were docked using the fast docking program Dock6.4. A grid-based scoring function that provides approximate molecular mechanics interaction energies consisting of van der Waals and electrostatic components was used to rank-order the compounds, and the top-ranked 1000 compounds were subjected to flexible molecular docking simulations using

RosettaLigand in the second stage. In this step, rank-ordering was performed using the RosettaLigand energy function, which includes terms for van der Waals attractive and repulsive forces, electrostatic interactions between pairs of amino acids, solvation assessing the effects of both side-chain–side-chain interactions and side-chain–ligand interactions, a statistical energy term derived from the probability of observing a side-chain conformation in the PDB, and an orientation-dependent hydrogen-bonding potential. The RosettaLigand energy function and protein–ligand interaction scores (InterfaceDelta term in the Rosetta energy function) were used to prepare rankings, and

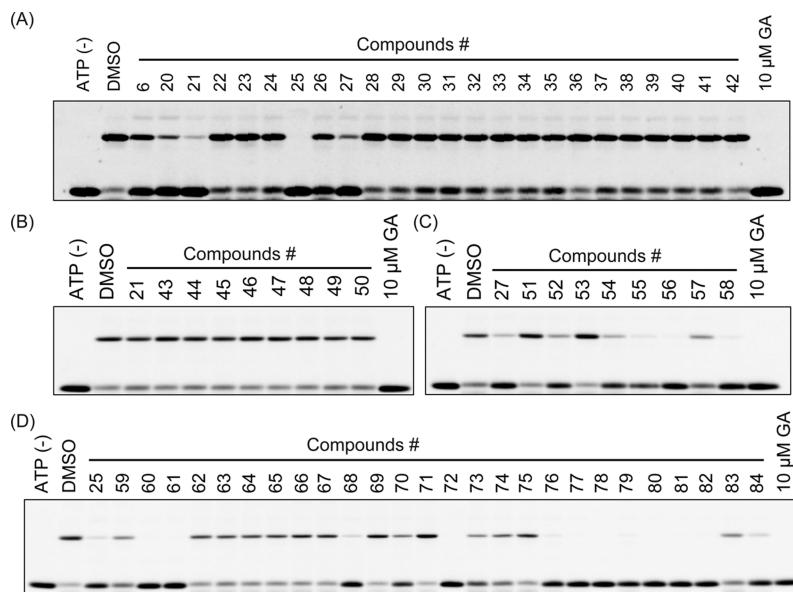


Figure 4. Inhibitory activities of commercially available analogues of (A) compound 6, (B) compound 21, (C) compound 27, and (D) compound 25 using an *in vitro* protein sumoylation assay.

the top-ranked 200 compounds were analyzed visually for formation of critical interactions. Finally, a set of 19 compounds were selected and acquired from commercial chemical vendors for biological assays.

Biological Validation of Virtually Screened Compounds. The ability of virtually screened compounds to inhibit Ubc9 was assessed using the *in vitro* sumoylation assay previously reported by Fukuda et al.⁵⁸ With this *in vitro* sumoylation assay, the 19 compounds selected after virtual screening were evaluated for their ability to inhibit sumoylation of RanGAP1-C2, a C-terminal fragment of protein substrate RanGAP1. The compounds were tested at a concentration of 200 μ M, and ginkgolic acid (GA), a known sumoylation inhibitor,⁵⁸ was used as a positive control. The results revealed that seven out of the 19 compounds (3, 4, 6, 10, 12, 18, and 19) inhibited at least 50% of RanGAP1 sumoylation (Figure 3A). Compound 6 was found to be the most active among the tested compounds, as it almost completely inhibited the sumoylation of RanGAP1 (Figure 3A).

As the *in vitro* sumoylation assay quantifies sumoylated RanGAP1, which requires the presence of SUMO E1 and Ubc9 only (not SUMO E3), the inhibition of Ubc9 by compounds 3, 4, 6, 10, 12, 18, and 19 was confirmed by detecting the SUMO E1-biotinylated SUMO-1 or SUMO E2-biotinylated SUMO-1 thioester intermediate. As shown in Figure 3B, the band corresponding to the SUMO E1-biotinylated SUMO-1 thioester intermediate was detected for compounds 3, 4, 6, 10, 12, and 18. This band was absent in case of compound 19, which indicates the inhibition of either SUMO E1 or Ubc9. To further confirm the Ubc9 inhibitory activity, compounds 3, 4, 6, 10, 12, 18, and 19 were then checked for blockage of the SUMO E2-biotinylated SUMO-1 thioester intermediate using the assay described previously.³⁰ The band corresponding SUMO E2-biotinylated SUMO-1 thioester intermediate could not be detected only in the case of compounds 4 and 6 (Figure 3C), confirming the Ubc9 inhibitory activity of compounds 4 and 6. The presence of bands corresponding to SUMO E2-biotinylated SUMO-1 thioester intermediate in the case of compounds 3, 10, 12, and 18 suggests that these compounds are either false positives or

may not inhibit sumoylation by blocking the formation of the SUMO E2-biotinylated SUMO-1 thioester intermediate.

Since compound 6 inhibited RanGAP1 sumoylation by blocking the formation of the SUMO E2-biotinylated SUMO-1 thioester intermediate (Figure 3), it was carried forward and evaluated in dose-dependent experiments by varying the concentration in the range from 0 to 400 μ M. The results of the assay confirmed that compound 6 inhibits Ubc9 in a concentration-dependent manner, and the IC₅₀ was found to be 265.3 μ M (Table 1). Although compound 6 displays very weak Ubc9 inhibitory activity, it represents a new scaffold that is significantly different from the other two classes of natural products reported previously.^{29,30} Compound 6 is (2R)-N-[4-[[2-(benzylcarbamoyl)acetyl]amino]phenyl]-2-(1,2,4-triazol-1-yl)propanamide, and it could be used as a starting point to identify more potent compounds by similarity searching. Consequently, compounds structurally similar to compound 6 were retrieved in order to improve the inhibitory activity. To retrieve similar compounds, two rounds of similarity searching were carried out against the ZINC database⁴¹ using MACCS structural keys.⁷⁰ In the first round, 23 compounds (20–42) were tested at a concentration of 200 μ M (Figure 4A), which disclosed three compounds with IC₅₀ values better than compound 6 (21, 25, and 27; Table 1).

Compounds 21, 25, and 27 share the N-phenyl-2-(1H-1,2,4-triazol-1-yl)propanamide scaffold with compound 6 and differ in chemical substitutions at the 4-position of the phenyl ring (Table 1). In the second round of similarity searching, compounds 21, 25, and 27 were used as query molecules to retrieve another set of 42 compounds (43–84) to be tested using the *in vitro* sumoylation assay. As shown in Figure 4B, no improvement in the activity was observed for the compound 21 analogues, whereas five analogues of compound 27 (52, 54, 55, 56, and 58) and 11 analogues of compound 25 (60, 61, 68, 72, 77–82, and 84) showed Ubc9 inhibitory activities at 200 μ M (Figure 4C,D). Overall, 65 analogues of compound 6 were tested using the *in vitro* protein sumoylation assay, out of which 20 were subjected to dose-dependent analysis. The compounds were tested at concentrations ranging from 0 to 400 μ M. The dose-dependent

Table 1. IC₅₀ Values of a Few Virtually Screened Compounds from an in Vitro Protein Sumoylation Assay

Compound	Structure	IC ₅₀ (μM)
6		265.3
20		333.1
21		199.2
25		68.8
27		210.8
55		157.7
56		213.3
60		46.3
61		160.2
72		262.1
82		167

experiments on these hits yielded nine compounds (**21**, **25**, **27**, **55**, **56**, **60**, **61**, **72**, and **82**) with IC₅₀ values lower than that of compound **6** (Table 1). Among these nine compounds, two compounds (**25** and **60**) displayed reasonable inhibitory activities of 68.8 and 46.3 μM respectively. Both compounds inhibited the sumoylation of RanGAP1 in a concentration-dependent manner.

Computational Analysis of the Binding Mode of Active Hits.

To understand the inhibitory mechanism of compounds **25**

and **60**, the docking-predicted binding mode was analyzed. As shown in Figure 5A,B and Figure S3 in the Supporting Information, both compounds are predicted to bind to Ubc9's predicted pocket in similar extended conformations. The binding of both compound **25** and compound **60** is driven mostly by hydrophobic interactions with Leu63, Phe64, Pro73, Tyr87, Val92, Cys93, Leu94, and Leu97. Most of these involve residues that are functionally important and found to be conserved, especially Pro73, Val92, Cys93, and Leu94 (Figure S2 in the Supporting Information). As far as hydrogen-bonding interactions are concerned, compound **25** forms only one hydrogen bond with the backbone carbonyl of Pro72. Compound **60** in the predicted binding mode instead forms several hydrogen bonds with Ubc9. These include hydrogen bonds with backbone of Leu94 and the side chains of Glu98 and Lys101. To shed further light on the binding mode of compound **25** and **60**, we generated and compared the pharmacophore features using the MOE program.⁴⁶ For the docking-predicted binding modes of compounds **25** and **60**, three hydrophobic features, one aromatic feature, one hydrogen-bond donor feature, and one hydrogen-bond acceptor feature could be overlaid without alignment (Figure S4 in the Supporting Information). All of the pharmacophore features lie in the receptor region favorable for their respective interactions. For example, the three hydrophobic features are in good positions to form hydrophobic contacts with Leu63, Phe64, Pro73, Tyr87, Val92, Cys93, Leu94, and Leu97. Moreover, the triazole and diazole rings in compounds **25** and **60**, respectively, are also in good positions to interact hydrophobically with the Phe64 and Lys65 side chains. Furthermore, the hydrogen-bond donor feature lies within hydrogen-bonding distance of the main acceptor atom of Pro72.

We next compared the docking-predicted binding modes of compounds **25** and **60** to that of spectomycin B1. Compounds **25** and **60** were predicted to bind in the same manner as spectomycin B1 (Figure 5 and Figure S3 in the Supporting Information). Spectomycin B1 is predicted to bind to Ubc9 with strong affinity and was found to form four hydrogen bonds with Ubc9 catalytic-site residues, including the main-chain amide of Leu94 and the side chains of Lys65, Glu98, and Asp127 (Figure S5C and Figure S3C in the Supporting Information). The binding of spectomycin B1 is also stabilized by the hydrophobic contacts between the spectomycin B1 aromatic core and the amino acid residues Pro72, Pro73, Val92, Leu94, and Leu97. Additionally, the hydrocarbon chains of Lys74, Glu98, Asn124, Gln126, and Asp127 also contribute to hydrophobic interactions (Figure S5C and Figure S3C in the Supporting Information). Although the two compounds make similar hydrophobic contacts with Ubc9, fewer hydrogen-bonding interactions were observed compared with spectomycin B1. Strong hydrogen-bonding interactions with Ubc9 may be one of the reasons for the higher activity of spectomycin B1 compared with compounds **25** and **60**. Therefore, it is expected that substitution of negatively charged functional groups may improve the potency of the identified scaffold.

To investigate ligand-induced changes in the binding site that were not apparent from the docking as well as to study the overall binding mode stability of compounds **25** and **60**, a 200 ns MD simulation was carried out using the Desmond program.⁵⁹ The average RMSDs of the protein backbone atoms from the starting frame were found to be 2.86 Å and 2.62 Å for compounds **25** and **60** respectively. No significant movements in the protein were observed in the case of the predicted Ubc9–compound **60**

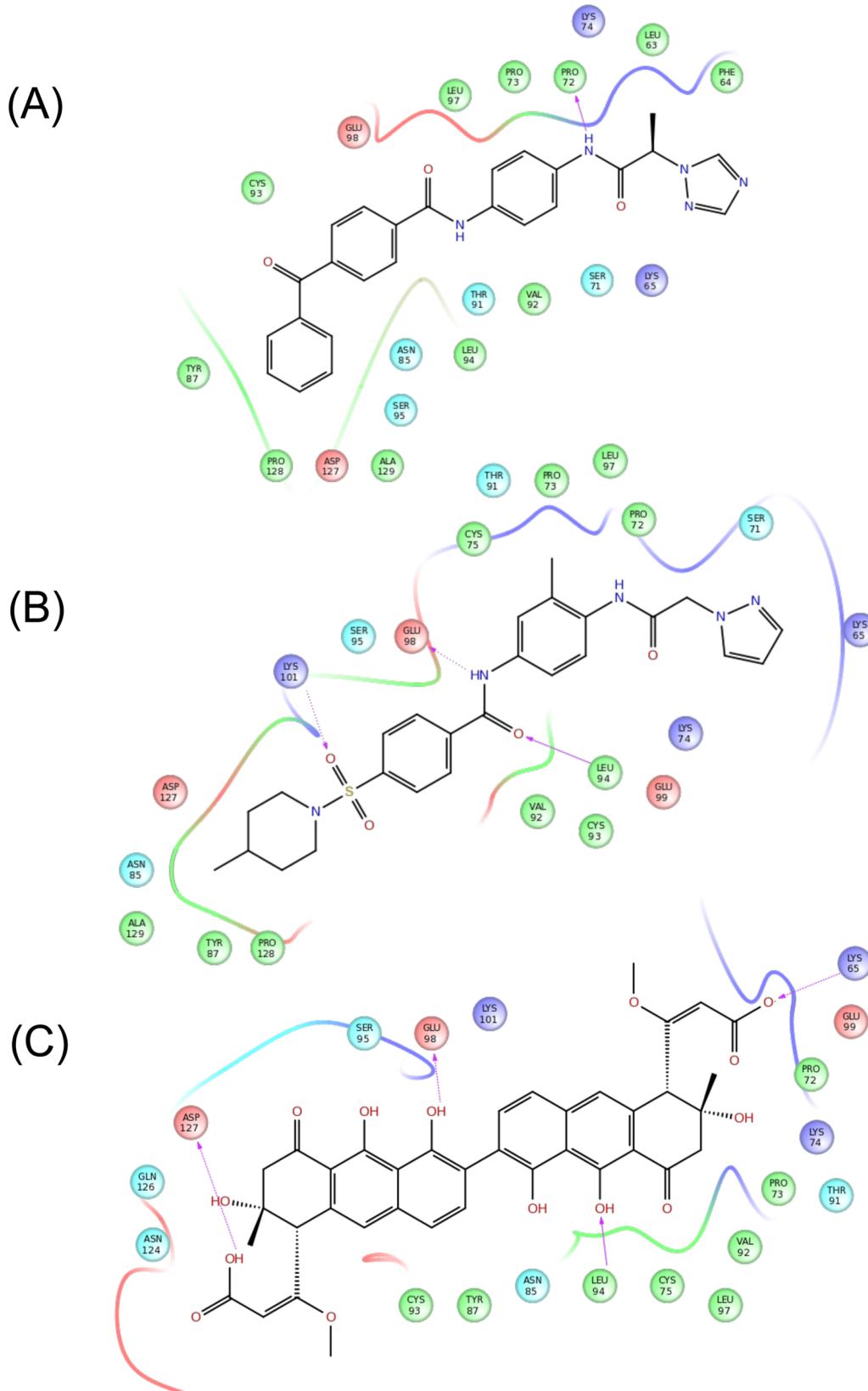


Figure 5. Docking-predicted binding modes of (A) compound 2S, (B) compound 60, and (C) spectomycin B1. Hydrogen bonds for the main chain and side chains are represented by solid and dotted magenta lines, respectively. Hydrophobic residues are colored in green, while negatively and positively charged residues are shown in red and blue, respectively. Polar residues are represented in cyan.

complex, which suggests a stable binding site and no ligand-induced conformational changes (Figure 6A). However, the

RMSD and root-mean-square fluctuation (RMSF) analyses (Figure 6A,B) highlighted possible protein movements involving

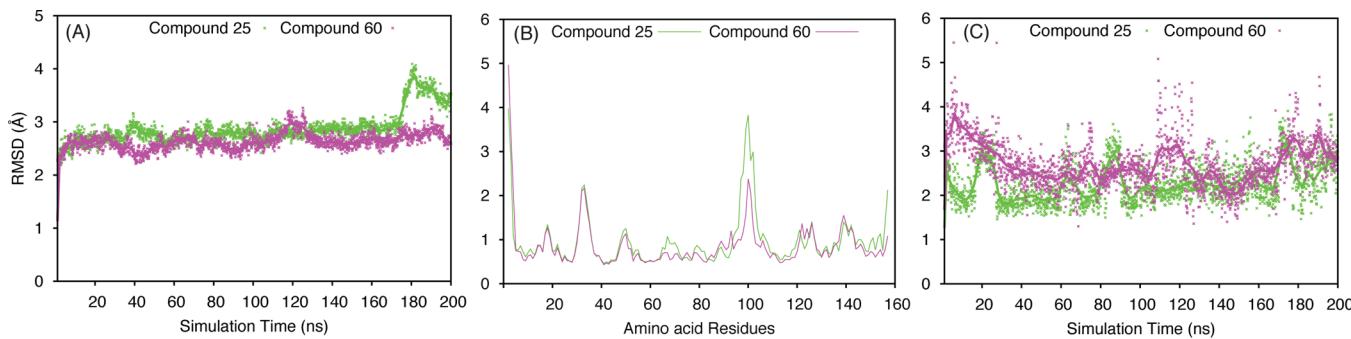


Figure 6. (A) Root-mean-square deviations and (B) root-mean-square fluctuations during 200 ns molecular dynamics trajectories for the docking-predicted protein–ligand complexes. (C) Root-mean-square deviations of the ligands.

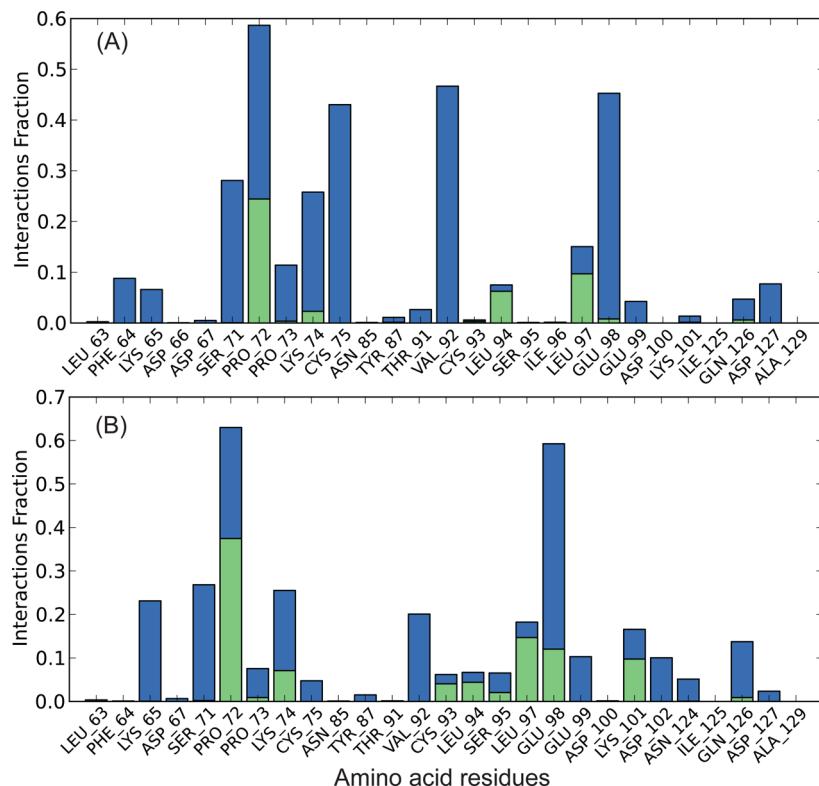


Figure 7. Hydrogen-bonding analysis using the molecular dynamics trajectories of (A) the docking-predicted Ubc9–compound 25 complex and (B) the docking-predicted Ubc9–compound 60 complex. For each complex, the green histogram represents direct hydrogen bonding between the ligand and the protein while the blue histogram represents water-mediated hydrogen bonding.

Glu97, Glu98, Glu99, Asp100, Lys101, and Asp102 for the predicted Ubc9–compound 25 complex. This movement was also evident from analysis of the ligand RMSDs, where the compound 25 movements are relatively larger than those of compound 60 (Figure 6C). The MD trajectories for compounds 25 and 60 were then analyzed for the presence of hydrogen bonds. Both direct and water-mediated hydrogen bonds were considered in the analysis, and the results are presented in Figure 7. The top-ranked docking-predicted pose for compound 25 displayed the formation of a hydrogen bond between the backbone carbonyl of Pro72 and the amide nitrogen near the triazole moiety. This hydrogen bond was maintained in the major part of the MD trajectory either as a direct or water-mediated hydrogen bond (Figure 7A). It is interesting to note that the MD simulations captured the formation of this hydrogen bond for compound 60 with high occupancy, as shown in Figure 7B. Additionally, the MD simulations further exposed three water-

mediated hydrogen bonds between compound 25 and Cys75, Val92, and Glu98 with more than 40% occupancy (Figure 7A). These hydrogen bonds were not previously detected because of the lack of explicit water molecules in the molecular docking calculations, which suggests the importance of MD simulations in identifying additional interactions. It should be further noted that only one of the three hydrogen bonds detected by molecular docking in the case of compound 60 (the hydrogen bond with Glu98) was retained (Figure 7B). Molecular docking of compounds 25 and 60 and spectomycin B1 and MD analysis of compounds 25 and 60 revealed binding preferences for Ubc9 inhibitors that can be used to guide the optimization of identified compounds for potency. These compounds already make significant hydrophobic contacts, and therefore, improvement in the hydrogen-bonding capabilities of the compounds should be considered as a strategy to design new compounds. The predicted binding pocket accommodates several positively and

negatively charged residues, such as Lys65, Lys74, Glu98, Lys101, and Asp127, and hence, the introduction of corresponding negatively and positively charged functional groups may improve the activity.

CONCLUSION

Structure-based virtual screening on the predicted small-molecule binding pocket on the surface of Ubc9 was performed. Similarity searching and molecular docking prioritized 19 compounds for evaluation of Ubc9 inhibitory activity. An *in vitro* sumoylation assay followed by a structural similarity search of active hits revealed two compounds with moderate Ubc9 inhibitory activities. The molecular docking and associated MD simulations suggest that compound binding may restrict access to residue Cys93, which catalyzes the formation of the E2–SUMO1 thioester intermediate. Although the compounds reported here are only moderately active, their structural novelty may make them good starting points for the development of new therapeutic agents against various diseases by targeting Ubc9.

ASSOCIATED CONTENT

Supporting Information

Human Ubc9 crystal structures used for pocket prediction with SiteMap predicted sites and properties (Table S1); top-ranked pocket in Ubc9 predicted by SiteMap and computational fragment mapping calculations on Ubc9 crystal structures by FTMap (Figure S1); sequence alignment of human Ubc9 with Ubc9 from other species and homologous proteins (Figure S2); docking-predicted binding modes of compounds **25** and **60** and spectomycin B1 (Figure S3); and pharmacophore features derived from docking-predicted binding modes of compounds **25** and **60** (Figure S4). This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: yoshidam@riken.jp (M.Y.).

*E-mail: kamzhang@riken.jp (K.Y.J.Z.).

Author Contributions

¹A.K. and A.I. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the RIKEN Integrated Cluster of Clusters (RICC) at RIKEN for the supercomputing resources used for this study. We thank members of our laboratories for help and discussions. This work was supported in part by the Initiative Research Unit Program and the Chemical Genomics Research Project, RIKEN, the CREST Research Project, the Japan Science and Technology Agency, and Grants-in-Aid for Scientific Research (C) and on Priority Area “Cancer” from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

- Baba, D.; Maita, N.; Jee, J. G.; Uchimura, Y.; Saitoh, H.; Sugawara, K.; Hanaoka, F.; Tochio, H.; Hiroaki, H.; Shirakawa, M. Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. *Nature* **2005**, *435*, 979–982.
- Geiss-Friedlander, R.; Melchior, F. Concepts in sumoylation: A decade on. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 947–956.
- Johnson, E. S. Protein modification by SUMO. *Annu. Rev. Biochem.* **2004**, *73*, 355–382.

(4) Palancade, B.; Doye, V. Sumoylating and desumoylating enzymes at nuclear pores: Underpinning their unexpected duties? *Trends Cell Biol.* **2008**, *18*, 174–183.

(5) Gareau, J. R.; Lima, C. D. The SUMO pathway: Emerging mechanisms that shape specificity, conjugation and recognition. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 861–871.

(6) Yeh, E. T. H. SUMOylation and De-SUMOylation: Wrestling with Life’s Processes. *J. Biol. Chem.* **2009**, *284*, 8223–8227.

(7) Zhao, J. Sumoylation regulates diverse biological processes. *Cell. Mol. Life Sci.* **2007**, *64*, 3017–3033.

(8) Capili, A. D.; Lima, C. D. Taking it step by step: Mechanistic insights from structural studies of ubiquitin/ubiquitin-like protein modification pathways. *Curr. Opin. Struct. Biol.* **2007**, *17*, 726–735.

(9) Wilkinson, K. A.; Henley, J. M., Mechanisms, regulation and consequences of protein SUMOylation. *Biochem. J.* **428**, 133–145.

(10) Wang, Y.; Dasso, M. SUMOylation and deSUMOylation at a glance. *J. Cell Sci.* **2009**, *122*, 4249–4252.

(11) Tong, H.; Hateboer, G.; Perrakis, A.; Bernards, R.; Sixma, T. K. Crystal Structure of Murine/Human Ubc9 Provides Insight into the Variability of the Ubiquitin-Conjugating System. *J. Biol. Chem.* **1997**, *272*, 21381–21387.

(12) Bernier-Villamor, V.; Sampson, D. A.; Matunis, M. J.; Lima, C. D. Structural Basis for E2-Mediated SUMO Conjugation Revealed by a Complex between Ubiquitin-Conjugating Enzyme Ubc9 and RanGAP1. *Cell* **2002**, *108*, 345–356.

(13) Wang, J.; Hu, W.; Cai, S.; Lee, B.; Song, J.; Chen, Y. The Intrinsic Affinity between E2 and the Cys Domain of E1 in Ubiquitin-like Modifications. *Mol. Cell* **2007**, *27*, 228–237.

(14) Mo, Y. Y.; Moschos, S. J. Targeting Ubc9 for cancer therapy. *Expert Opin. Ther. Targets* **2005**, *9*, 1203–1216.

(15) Mo, Y. Y.; Yu, Y.; Theodosiou, E.; Ee, P. L.; Beck, W. T. A role for Ubc9 in tumorigenesis. *Oncogene* **2005**, *24*, 2677–2683.

(16) Kim, K. I.; Baek, S. H. SUMOylation Code in Cancer Development and Metastasis. *Mol. Cells* **2006**, *22*, 247–253.

(17) Sarge, K. D.; Park-Sarge, O. K. Sumoylation and human disease pathogenesis. *Trends Biochem. Sci.* **2009**, *34*, 200–205.

(18) Duan, X.; Trent, J. O.; Ye, H. Targeting the SUMO E2 conjugating enzyme Ubc9 interaction for anti-cancer drug design. *Anticancer Agents Med. Chem.* **2009**, *9*, 51–54.

(19) Driscoll, J.; DeChowdhury, R. Therapeutically targeting the SUMOylation, ubiquitination and proteasome pathways as a novel anticancer strategy. *Targeted Oncol.* **2010**, *5*, 281–289.

(20) Sarge, K. D.; Park-Sarge, O.-K. SUMO and Its Role in Human Diseases. *Int. Rev. Cell Mol. Biol.* **2011**, *288*, 167–183.

(21) Voutsadakis, I. Ubiquitin- and ubiquitin-like proteins-conjugating enzymes (E2s) in breast cancer. *Mol. Biol. Rep.* **2013**, *40*, 2019–2034.

(22) Dong, M.; Pang, X.; Xu, Y.; Wen, F.; Zhang, Y. Ubiquitin-Conjugating Enzyme 9 Promotes Epithelial Ovarian Cancer Cell Proliferation in Vitro. *Int. J. Mol. Sci.* **2013**, *14*, 11061–11071.

(23) Moschos, S. J.; Smith, A. P.; Mandic, M.; Athanassiou, C.; Watson-Hurst, K.; Jukic, D. M.; Edington, H. D.; Kirkwood, J. M.; Becker, D. SAGE and antibody array analysis of melanoma-infiltrated lymph nodes: Identification of Ubc9 as an important molecule in advanced-stage melanomas. *Oncogene* **2007**, *26*, 4216–4225.

(24) Moschos, S. J.; Jukic, D. M.; Athanassiou, C.; Bhargava, R.; Dacic, S.; Wang, X.; Kuan, S.-F.; Fayewicz, S. L.; Galambos, C.; Acquafonda, M.; Dhir, R.; Becker, D. Expression analysis of Ubc9, the single small ubiquitin-like modifier (SUMO) E2 conjugating enzyme, in normal and malignant tissues. *Hum. Pathol.* **2010**, *41*, 1286–1298.

(25) McDoneli-Silvers, A. L.; Nimri, C. F.; Stoner, G. D.; Lubet, R. A.; You, M. Differential gene expression in human lung adenocarcinomas and squamous cell carcinomas. *Clin. Cancer. Res.* **2002**, *8*, 1127–1138.

(26) Gocke, C. B.; Yu, H.; Kang, J. Systematic identification and analysis of mammalian small ubiquitin-like modifier substrates. *J. Biol. Chem.* **2005**, *280*, 5004–5012.

(27) Seeler, J. S.; Bischof, O.; Nacerddine, K.; Dejean, A. SUMO, the three Rs and cancer. *Curr. Top. Microbiol. Immunol.* **2007**, *313*, 49–71.

(28) Zhong, N.; Radu, G.; Ju, W.; Brown, W. T. Novel progerin-interacting partner proteins hnRNP E1, EGF, Mel 18, and UBC9

- interact with lamin A/C. *Biochem. Biophys. Res. Commun.* **2005**, *338*, 855–861.
- (29) Kim, Y. S.; Nagy, K.; Keyser, S.; Schneekloth, J. S., Jr. An Electrophoretic Mobility Shift Assay Identifies a Mechanistically Unique Inhibitor of Protein Sumoylation. *Chem. Biol.* **2013**, *20*, 604–613.
- (30) 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, K. Y. J.; Yoshida, M.; Ito, A. Spectomycin B1 as a Novel SUMOylation Inhibitor That Directly Binds to SUMO E2. *ACS Chem. Biol.* **2013**, *8*, 2635–2642.
- (31) Kim, Y. S.; Keyser, S. G. L.; Schneekloth, J. S., Jr. Synthesis of 2',3',4'-trihydroxyflavone (2-D08), an inhibitor of protein sumoylation. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 1094–1097.
- (32) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* **1977**, *112*, 535–542.
- (33) Maestro, version 9.2; Schrödinger, LLC: New York, 2011.
- (34) SiteMap, version 2.5; Schrödinger, LLC: New York, 2011.
- (35) Halgren, T. New Method for Fast and Accurate Binding-Site Identification and Analysis. *Chem. Biol. Drug Des.* **2007**, *69*, 146–148.
- (36) Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377–389.
- (37) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
- (38) The PyMOL Molecular Graphics System, version 1.5.0.4; Schrödinger, LLC: New York, 2010.
- (39) Williams, T.; Kelley, C. Gnuplot 4.6: An Interactive Plotting Program, 2012.
- (40) Vainio, M. J.; Puranen, J. S.; Johnson, M. S. ShaEP: Molecular Overlay Based on Shape and Electrostatic Potential. *J. Chem. Inf. Model.* **2009**, *49*, 492–502.
- (41) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A Free Tool To Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52*, 1757–1768.
- (42) OMEGA, version 2.4.6; OpenEye Scientific Software, Inc.: Santa Fe, NM, 2012; <http://www.eyesopen.com>.
- (43) Hawkins, P. C.; Nicholls, A. Conformer generation with OMEGA: Learning from the data set and the analysis of failures. *J. Chem. Inf. Model.* **2012**, *52*, 2919–2936.
- (44) Hawkins, P. C.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; Stahl, M. T. Conformer generation with OMEGA: Algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **2010**, *50*, 572–584.
- (45) Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490–519.
- (46) Molecular Operating Environment (MOE), version 2011.10; Chemical Computing Group Inc.: Montreal, QC, Canada, 2010.
- (47) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J. Comput. Chem.* **2000**, *21*, 132–146.
- (48) Labute, P. Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins: Struct., Funct., Bioinf.* **2009**, *75*, 187–205.
- (49) Davis, I. W.; Baker, D. RosettaLigand docking with full ligand and receptor flexibility. *J. Mol. Biol.* **2009**, *385*, 381–392.
- (50) Davis, I. W.; Raha, K.; Head, M. S.; Baker, D. Blind docking of pharmaceutically relevant compounds using RosettaLigand. *Protein Sci.* **2009**, *18*, 1998–2002.
- (51) Meiler, J.; Baker, D. ROSETTALIGAND: Protein–small molecule docking with full side-chain flexibility. *Proteins* **2006**, *65*, 538–548.
- (52) Leaver-Fay, A.; Tyka, M.; Lewis, S. M.; Lange, O. F.; Thompson, J.; Jacak, R.; Kaufman, K. W.; Renfrew, P. D.; Smith, C. A.; Sheffler, W.; Davis, I. W.; Cooper, S.; Treuille, A.; Mandell, D. J.; Richter, F.; Ban, Y.-E. A.; Fleishman, S. J.; Corn, J. E.; Kim, D. E.; Lyskov, S.; Berrondo, M.; Mentzer, S.; Popović, Z.; Havranek, J. J.; Karanicolas, J.; Das, R.; Meiler, J.; Kortemme, T.; Gray, J. J.; Kuhlman, B.; Baker, D.; Bradley, P. Chapter nineteen—Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. *Methods Enzymol.* **2011**, *487*, 545–574.
- (53) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecule–ligand interactions. *J. Mol. Biol.* **1982**, *161*, 269–288.
- (54) Ewing, T. A.; Makino, S.; Skillman, A. G.; Kuntz, I. DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 411–428.
- (55) Moustakas, D.; Lang, P. T.; Pegg, S.; Pettersen, E.; Kuntz, I.; Brooijmans, N.; Rizzo, R. Development and validation of a modular, extensible docking program: DOCK 5. *J. Comput.-Aided Mol. Des.* **2006**, *20*, 601–619.
- (56) Lorber, D. M.; Shoichet, B. K. Flexible ligand docking using conformational ensembles. *Protein Sci.* **1998**, *7*, 938–950.
- (57) Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. Automated docking with grid-based energy evaluation. *J. Comput. Chem.* **1992**, *13*, 505–524.
- (58) Fukuda, I.; Ito, A.; Hirai, G.; Nishimura, S.; Kawasaki, H.; Saitoh, H.; Kimura, K.; Sodeoka, M.; Yoshida, M. Ginkgolic acid inhibits protein SUMOylation by blocking formation of the E1–SUMO intermediate. *Chem. Biol.* **2009**, *16*, 133–140.
- (59) Desmond Molecular Dynamics System, version 3.6; D. E. Shaw Research: New York, 2013. Maestro–Desmond Interoperability Tools, version 3.6; Schrödinger, LLC: New York, 2013.
- (60) Ryckaert, J.-P.; Cicotti, G.; Berendsen, H. J. C. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (61) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An *N*–log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (62) Campagna-Slater, V.; Mok, M. W.; Nguyen, K. T.; Feher, M.; Najmanovich, R.; Schapira, M. Structural Chemistry of the Histone Methyltransferases Cofactor Binding Site. *J. Chem. Inf. Model.* **2011**, *51*, 612–623.
- (63) Vidler, L. R.; Brown, N.; Knapp, S.; Hoelder, S. Druggability Analysis and Structural Classification of Bromodomain Acetyl-lysine Binding Sites. *J. Med. Chem.* **2012**, *55*, 7346–7359.
- (64) Yunus, A. A.; Lima, C. D. Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway. *Nat. Struct. Mol. Biol.* **2006**, *13*, 491–499.
- (65) Brenke, R.; Kozakov, D.; Chuang, G.-Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S. Fragment-based identification of druggable “hot spots” of proteins using Fourier domain correlation techniques. *Bioinformatics* **2009**, *25*, 621–627.
- (66) Rogers, D. J.; Tanimoto, T. T. A Computer Program for Classifying Plants. *Science* **1960**, *132*, 1115–1118.
- (67) Kumar, A.; Ito, A.; Hirohama, M.; Yoshida, M.; Zhang, K. Y. J. Identification of Sumoylation Activating Enzyme 1 Inhibitors by Structure-Based Virtual Screening. *J. Chem. Inf. Model.* **2013**, *53*, 809–820.
- (68) Kumar, A.; Ito, A.; Hirohama, M.; Yoshida, M.; Zhang, K. Y. J. Identification of quinazolinylxyloxy biaryl urea as a new class of SUMO activating enzyme 1 inhibitors. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 5145–5149.
- (69) Kumar, A.; Zhang, K. Y. J. Hierarchical virtual screening approaches in small molecule drug discovery. *Methods* **2014**, DOI: 10.1016/j.ymeth.2014.07.007.
- (70) MACCS Keys; MDL Information Systems, Inc.: San Leandro, CA.