Electronic Supplementary Material (ESI) for MedChemComm.
This journal is © The Royal Society of Chemistry 2014

Discovery of small molecule inhibitors targeting the SUMO-SIM interaction using a protein interface consensus approach

Electronic Supplementary Information

Arnout R.D. Voet,[‡a] Akihiro Ito,[‡bc] Mikako Hirohama,[be] Seiji Matsuoka,[d] Naoya Tochio,[§f] Takanori Kigawa,[f] Minoru Yoshida[bcde] and Kam Y.J. Zhang[*a]

[a] Zhang Initiative Research Unit, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. E-mail: kamzhang@riken.jp

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

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

[d] Drug Discovery Platforms Cooperation Division, RIKEN Center for Sustainable Resource Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan.

[e] Japan Science and Technology Corporation, CREST Research Project, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan.

[f] Laboratory for Biomolecular Structure and Dynamics, Quantitative Biology Center, RIKEN, 1-7-22 Suehiro, Yokohama, Kanagawa 230-0045, Japan.

[‡] These authors contributed equally to this work.

[§] Current Address: Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, 1-3-1 Kagamiyama, Hiroshima 739-8530, Japan.

Virtual screening protocol

Druggability analysis of the SUMO structure was performed using the HotPatch server[1], indicating an absence of a clear small molecule binding cleft, however a small hydrophobic trench is present. Electrostatic analysis of the SUMO-SIM interface was performed using APBS.[2]

The compound database (Namiki Shoji CO., LTD, Tokyo, Japan) used for virtual screening was prepared by enumerating the different stereo and tautomers using OpenEye's Quacpac and Flipper tools (OpenEye, Santa Fe, NM, USA). A maximum of 250 conformations for each unique molecule was generated using OpenEye's Omega2 tool.[3, 4] Using the Filter tool, molecules with a net positive charge or more hydrogen bond donors than acceptors were removed.

The consensus pharmacophore was created using MOE (Molecular Operating Environment, Chemical Computing Group, Montreal, Canada) pharmacophore implementation with the "PCHD scheme", based on SUMO-SIM interactions observed in the following PDBs: 1WM3, 1WYW, 1Z5S, 2ASQ, 2KQS, 3KYC and 3RD2. First the structures were superposed on the backbone of the SUMO proteins, and then the consensus interactions were computed for the SUMO interacting amino acids (and water molecules). The final pharmacophore query consisted of 3 "hydrophobic or aromatic" features (representing hydrophobic side chain interactions) and 2 hydrogen bond donor and 2 hydrogen bond acceptor features (including directionality of the interaction) that resemble the main chain interactions at the SUMO interface (supplementary Figure 1). As a forbidden volume, the union of the side chain conformations of PDB 1WYW and 1Z5S was chosen. During virtual screening using the pharmacophore query, a partial match was allowed, while the 3 hydrophobic/aromatic features and the 2 central hydrogen bonds were essential, it was sufficient to hit only one out of the 2 "outside" hydrogen bond features.

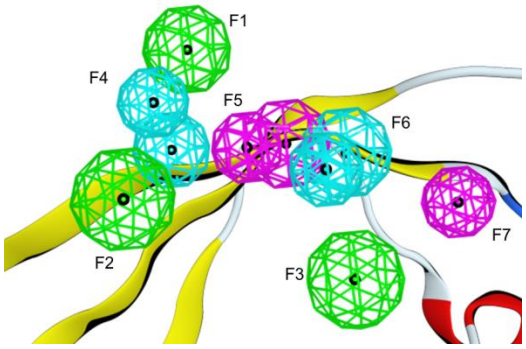


Figure S1: Graphical depiction of the pharmacophore query representing the key interactions at the SUMO-SIM interface. The hydrophobic/aromatic features (F1, F2 and F3) are presented as green spheres. The hydrogen bonds are represented as magenta or blue spheres for hydrogen bond acceptors or donors respectively. During the pharmacophore search, partial matching was allowed and it was sufficient to hit only one of the non-essential features F4 and F7.

After the pharmacophore query, the remaining compounds were docked using GOLD in virtual screening default settings[5], with the PLP scoring function[6], and for 10 runs per compound, to the 2 different receptor conformations. The docked conformations were rescored using DrugScoreX[7, 8]. Compounds that ranked in the 2 different receptors for the 2 different scoring functions in the lowest 50 percent were removed. The remaining compound binding modes were post-filtered using the pharmacophore query to remove undesired binding modes. Finally the compound binding modes were compared between the 2 different docking receptors and the compounds were retained only if the RMSD of the two binding modes for each compound was within 0.5 Å (absolute positions). In a subsequent step, the remaining binding modes were filtered using EleKit[9] for electrostatic similarity with the SIM DAXX peptide, removing all docked compounds without similarity. Finally the remaining compounds were selected after visual inspection of the binding mode to the different receptor conformations with the DrugScoreX[7, 8] visualization script for Pymol (Schrödinger LCC). Up to 3 derivatives belonging to the same chemotype were allowed to be selected.

After the first round, derivatives of the most promising active compounds (SSI-012, SSI-020, SSI-038 and SSI-062 belong to one chemical class, while SSI-041 and SSI-060 belong to other classes) were identified by similarity searching using MACCS and 3 point pharmacophore fingerprints using MOE.

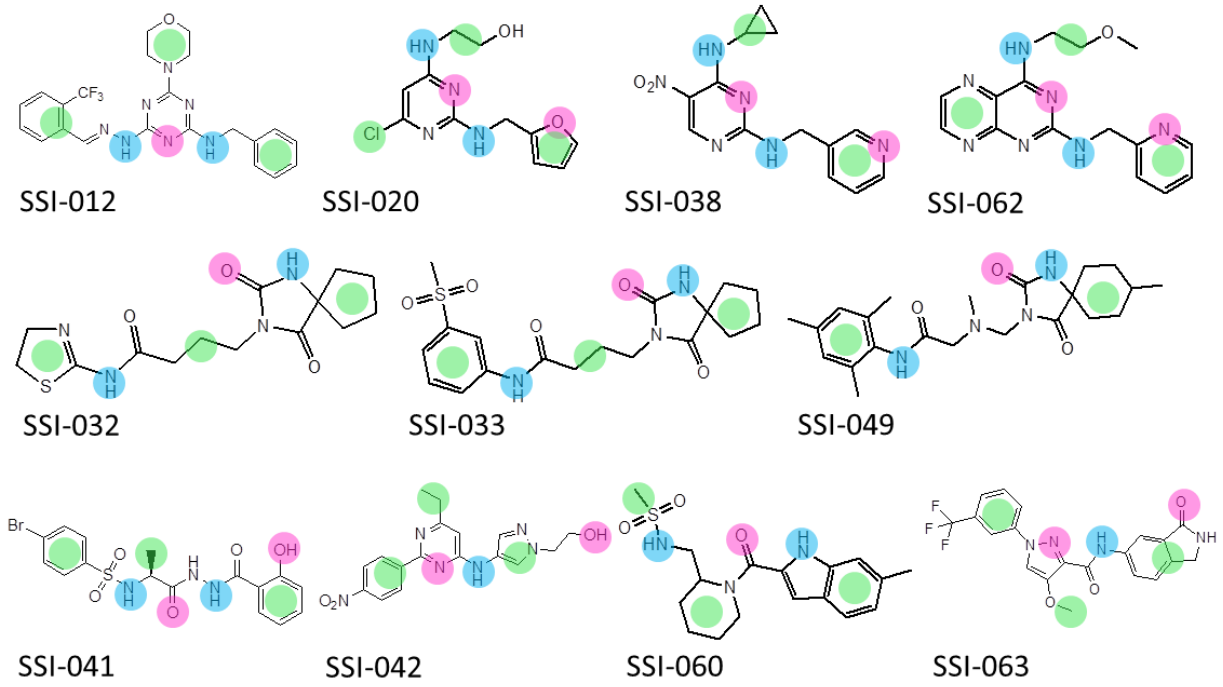


Figure S2: The first hit molecules oriented according to the pharmacophore query represented in supplementary Figure S1. Green spheres indicate the mapping of the hydrophobic/aromatic pharmacophore features, blue and magenta of the hydrogen bond donor/acceptor features respectively.

Bacterial protein expression and purification

Recombinant His-tagged SUMO1 protein was purified as described[10] with minor modifications. Briefly, the pET-based bacterial expression plasmids containing SUMO1 cDNA were introduced into Escherichia coli BL21 (DE3). The expression of recombinant proteins was induced by addition of 0.3 mM isopropyl-b-D-galactopyranoside (IPTG) for 8 h. Purification of His-tagged SUMO1 proteins were carried out using (Ni^{2+})-affinity column (QIAGEN). The SUMO2/3 isoform protein was purified similarly.

AlphaLISA assay

His- tagged SUMO1 (3 µM) and biotinylated DAXX-SIM peptide were incubated in 10 µl of a reaction buffer containing 20 mM HEPES (pH 7.5), 150 mM KCl, and 0.05% Tween20 at room temperature for 1 h. Then, nickel chelate acceptor beads and streptavidin-coated donor beads (20 µg/ml each) were sequentially added into the reaction and incubated in the dark at room temperature for additional 30 min, respectively. The signal was measured using an EnSpire™ (PerkinElmer).[11]

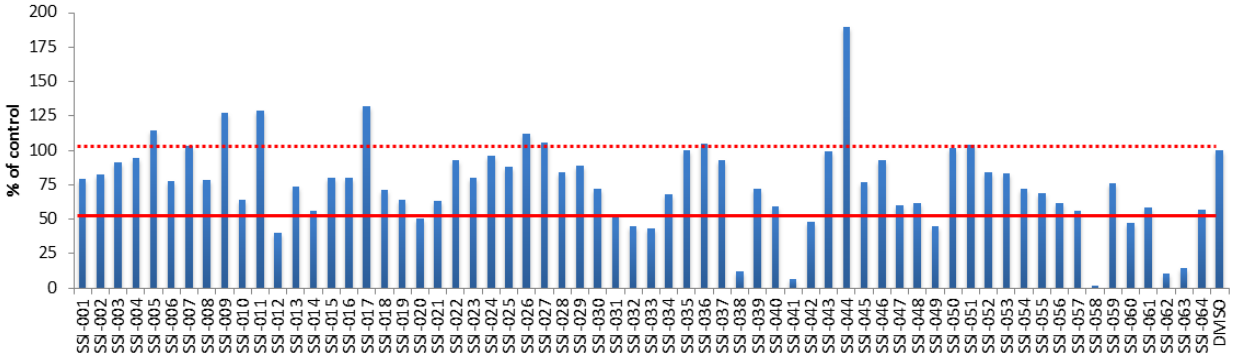


Figure S3: AlphaLISA results of the first round of compounds at 500µM concentration. The dotted and solid red lines indicate the 100% effect and 50% effect relative DMSO respectively.

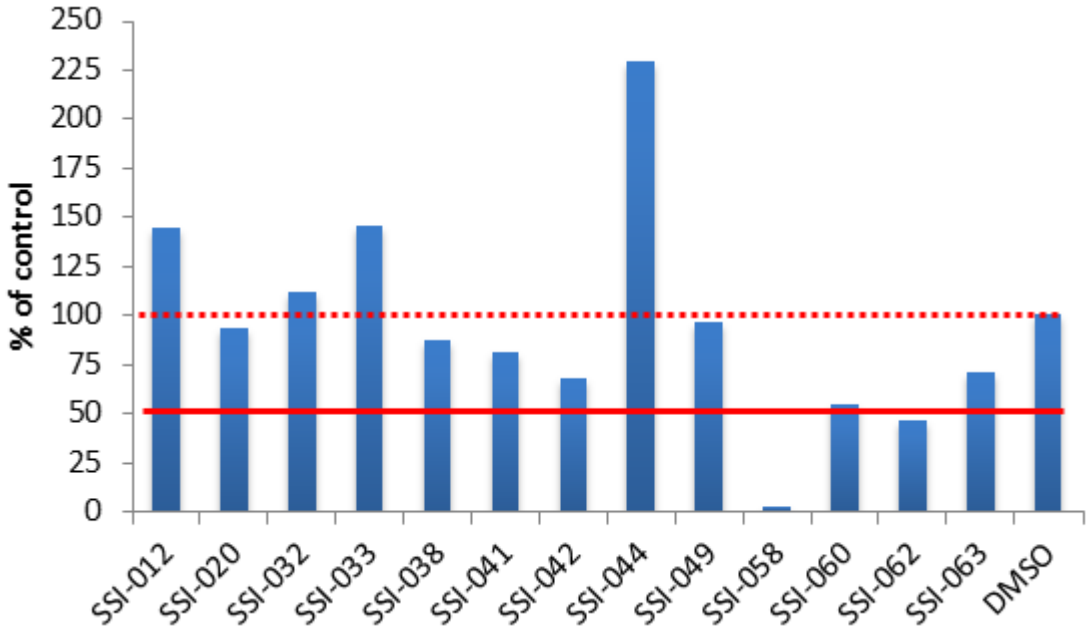


Figure S4: AlphaLISA results of the first round 500 µM hit compounds at 100µM compound concentration. The dotted and solid red lines indicate the 100% effect and 50% effect relative DMSO respectively.

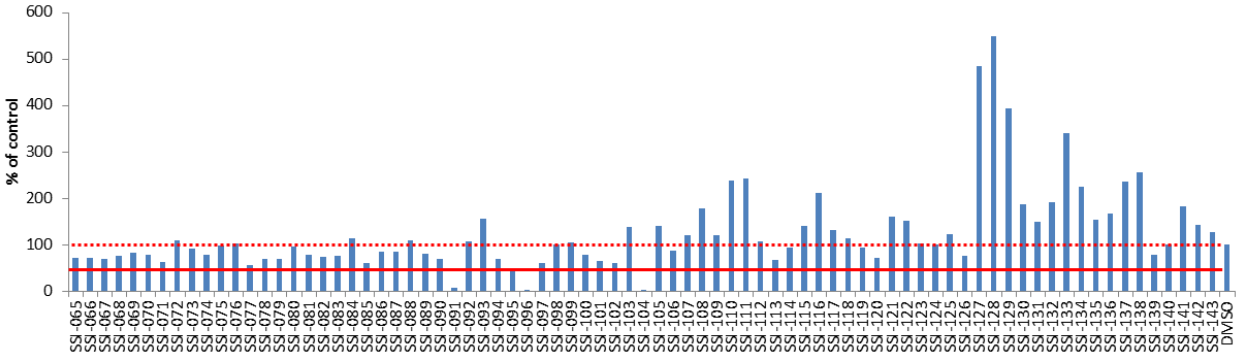


Figure S5: AlphaLISA results of the second round of compounds at 100µM concentration. The dotted and full red lines indicate the 100% effect and 50% effect relative DMSO respectively.

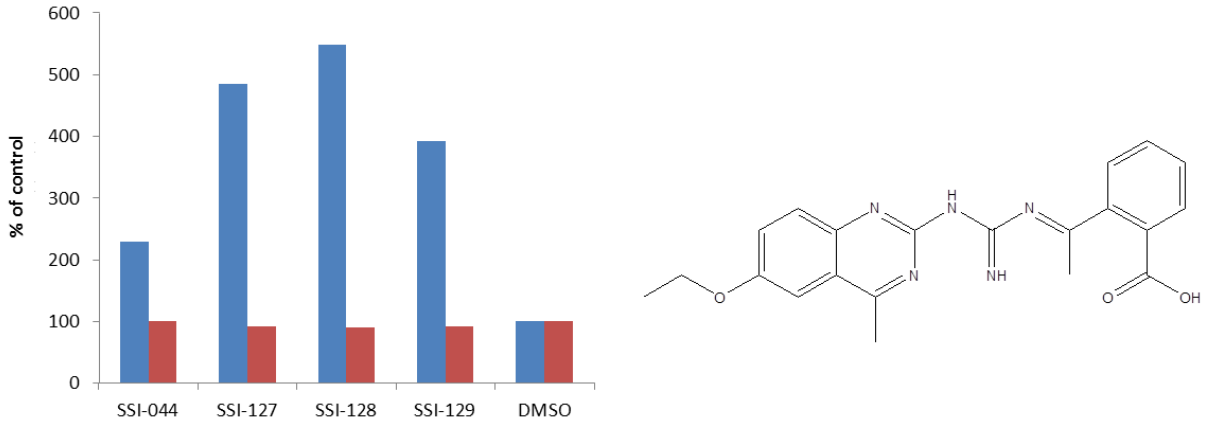


Figure S6: AlphaLISA results the SUMO-SIM interaction stimulators at 100µM compound concentration (blue) with the TruHits screen (red) indicating the stimulatory effect is not an artifact inherent to the compound structures. All chemical structures are derivatives of the same chemotype. On the right the chemical structure of the strongest stimulators SSI-128 is depicted.

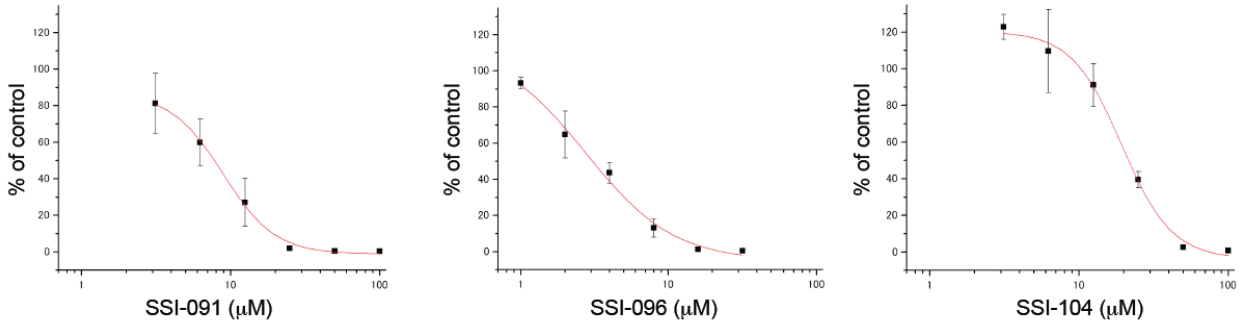


Figure S7: Dose response curves of the most potent compounds.

SPR binding Analysis

The SPR experiment was performed with a Biacore T200 (GE Healthcare). His-SUMO1 (or His-SUMO2/3) (0.46 µg/ml) in running buffer (HBS-P+, GE Healthcare) was immobilized to approximately 280 RU on a Series S Sensor Chip NTA (GE Healthcare) using a standard ligand capture method with a His Capture Kit (GE Healthcare). A flow cell without immobilized proteins was used as reference. The analyte solution was prepared by a serial dilution of compounds with the running buffer containing 5% DMSO. Binding analysis was conducted at a flow rate of 30 µl/min at 25°C. In each run, the association phase and the following dissociation phase were monitored for 60 and 120 seconds, respectively. Blank control run was performed by the injection of running buffer containing 5% DMSO. From the obtained reference-subtracted

sensorgrams, the dissociation constant (K_D) of compounds was estimated by a global fitting to a simple 1:1 binding model in the Biacore evaluation software (GE Healthcare). The sensorgrams were corrected for DMSO bulk response by using calibration curves obtained with running buffer containing 4-6% DMSO.

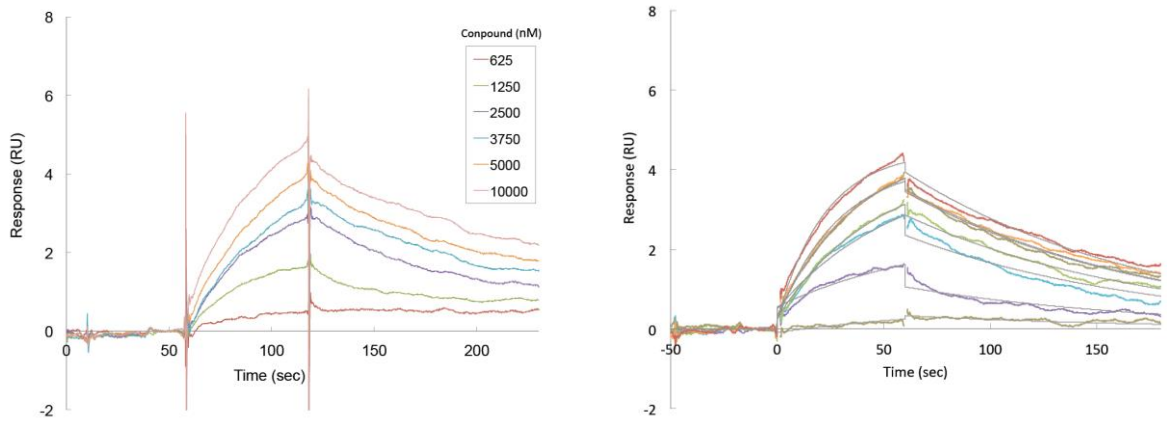


Figure S8: SPR sensorgrams of SSI-091 binding to immobilized SUMO1 (left). The sensorgrams were fitted to a 1:1 binding model by an analysis using the Biacore evaluation software (GE Healthcare) (right).

Table S1: Kinetic parameters of most potent compounds in SPR experiments

ID-number	Kon (M/S)		Koff (1/S)	
	SUMO-1	SUMO-2/3	SUMO-1	SUMO-2/3
SSI-091	3.6×10^{3}	4.0×10^{3}	6.4×10^{-3}	7.1×10^{-3}
SSI-096	1.5×10^{3}	1.0×10^{3}	6.9×10^{-2}	6.0×10^{-2}
SSI-104	5.2×10^{2}	1.3×10^{3}	1.8×10^{-2}	4.2×10^{-2}

Cellular split luciferase assay

The cellular split luciferase assay was performed using two mammalian expression constructs consisting of the N-terminal fragment of *Renilla* luciferase fused to SUMO1 and the C-terminal fragment of *Renilla* luciferase fused to the SIM sequence from the Daxx protein, respectively. 293T cells were transfected with the two above expression vetors, followed by treatment with inhibitors for 24 h. After cells were lysed in passive lysis buffer, the bioluminescence generated by adding coelenterazine into the cell lysates was measured using a luminometer.

Evaluated compounds

All evaluated compounds were purchased from Enamine or Vitas-M Laboratory through Namiki Shoji CO., LTD (Tokyo, Japan). A complete list of compounds with their activity and identity is shown at the end of this document

NMR experiments

All NMR spectra were recorded at 298 K on a 700 MHz Bruker Avance spectrometer equipped with CryoProbe.

The STD experiments were performed using non-labeled SUMO1 protein, purified as described above. The 1D ^{1}H NMR spectrum of 500 μM SSI-091 was measured in the presence of 12.3 μM non-labeled SUMO1 in 20 mM d_{18}-HEPES (pH7.5), 150 mM KCl, 5% d_6-DMSO, and 95% D_2O. The double off-resonance spectra were obtained when the mixture of SSI-091 and SUMO1 was irradiated at 40 ppm. The STD spectra were obtained by subtraction of the on-resonance spectrum (at 0.07 ppm) from the off-resonance spectrum (at 40 ppm).

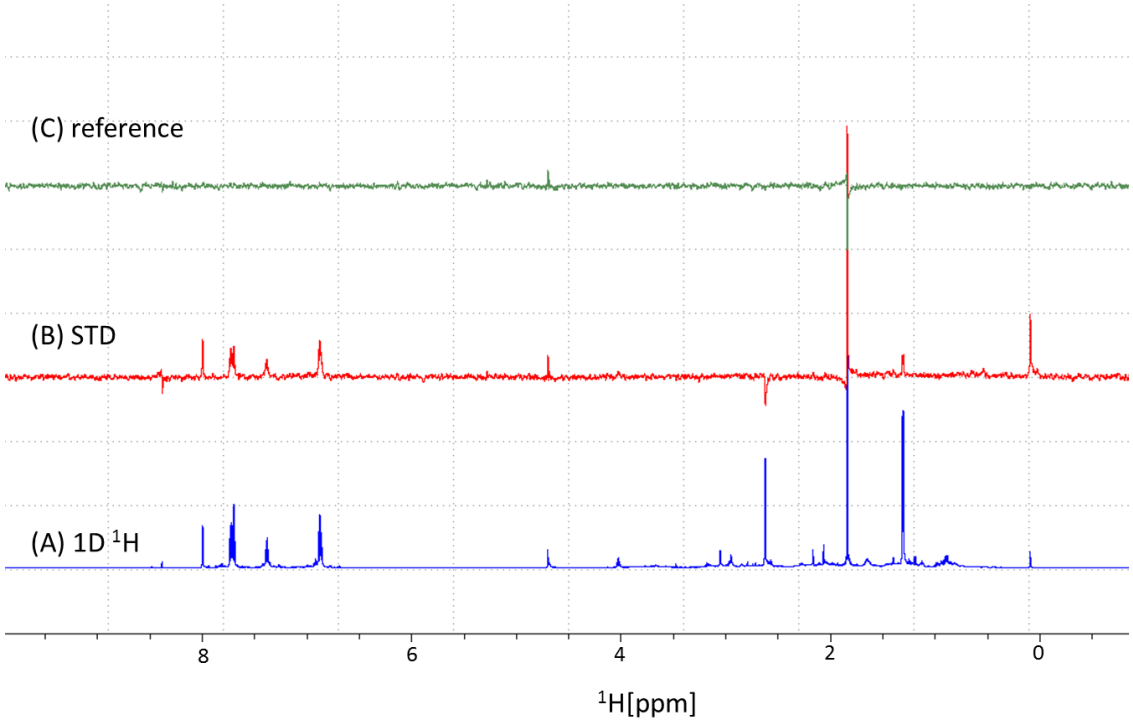
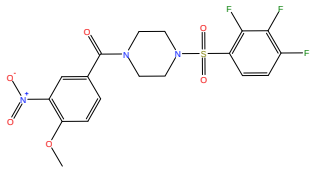


Figure S9: STD experiment for SSI-091.

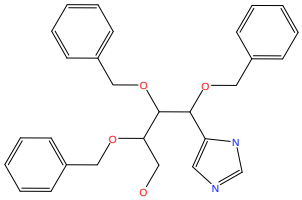
(A) 1D ^{1}H NMR spectrum of SSI-091 in the presence non-labeled SUMO1. (B) The STD spectra obtained by subtraction of the on-resonance spectrum (at 0.07 ppm) from the off-resonance spectrum (at 40 ppm). (C) The double off-resonance spectra of the mixture of SSI-091 and SUMO1.

The HSQC experiments were performed using ^{13}C, ^{15}N labeled SUMO1, produced in a cell-free expression system.[12-14] The backbone chemical shift assignments of SUMO1 were accomplished by using conventional 3D triple resonance spectra at concentration of about 1 mM SUMO1. All spectra were processed using NMRPipe[15], and the programs Kujira[16] and NMRView[17] were employed for optimal visualization and spectral analyses. To experimentally map the binding interface, the ^{1}H,^{15}N HSQC spectra were measured using 50 μM ^{13}C,^{15}N labeled SUMO1 (apo spectra) and 50 μM ^{13}C,^{15}N labeled SUMO1 with 500 μM SSI-091 in 20 mM d_{18}-HEPES (pH7.5), 150 mM KCl, 5% d_6-DMSO (holospectrum).

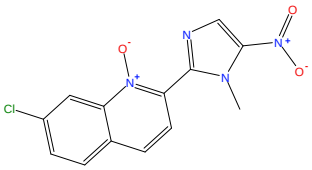
Table S2: List of structures of all evaluated compounds with percentage of inhibition measured by AlphaLisa



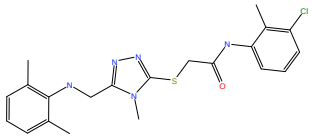
SSI-001
inhibition (%): 78.70



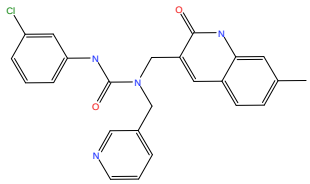
SSI-002
inhibition (%): 82.30



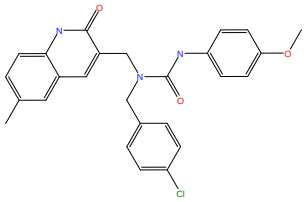
SSI-003
inhibition (%): 90.60



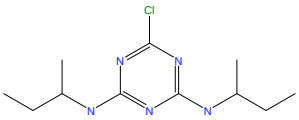
SSI-004
inhibition (%): 94.00



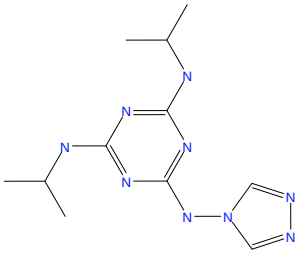
SSI-005
inhibition (%): 114.00



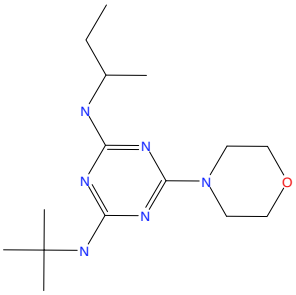
SSI-006
inhibition (%): 77.60



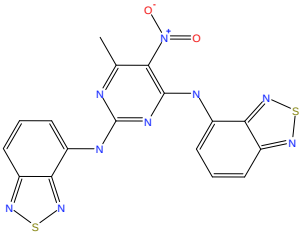
SSI-007
inhibition (%): 102.60



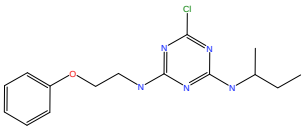
SSI-008
inhibition (%): 78.00



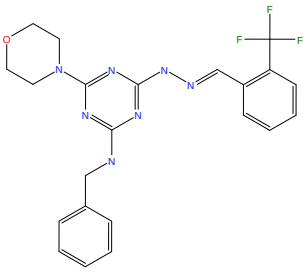
SSI-009
inhibition (%): 126.70



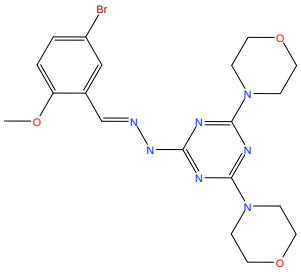
SSI-010
inhibition (%): 63.30



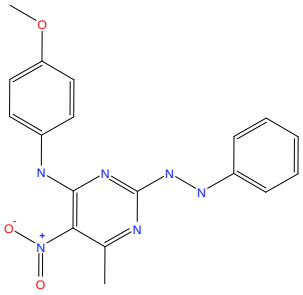
SSI-011
inhibition (%): 128.20



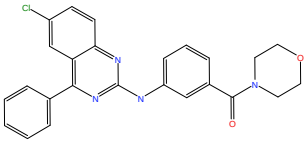
SSI-012
inhibition (%): 39.70



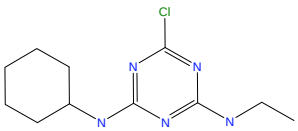
SSI-013
inhibition (%): 73.40



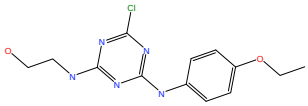
SSI-014
inhibition (%): 56.00



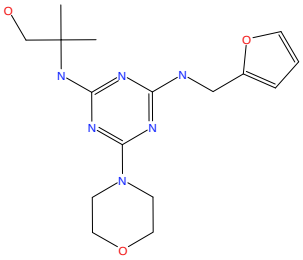
SSI-015
inhibition (%): 79.70



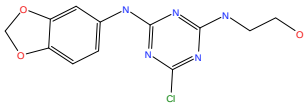
SSI-016
inhibition (%): 79.70



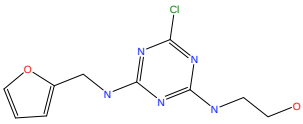
SSI-017
inhibition (%): 131.30



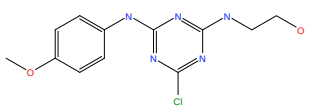
SSI-018
inhibition (%): 70.90



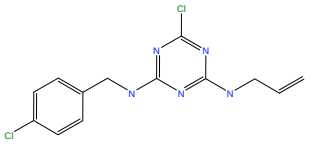
SSI-019
inhibition (%): 63.90



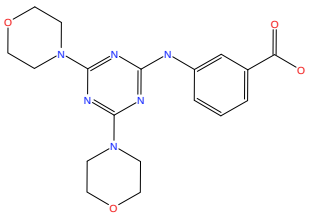
SSI-020
inhibition (%): 50.40



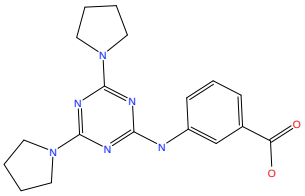
SSI-021
inhibition (%): 62.70



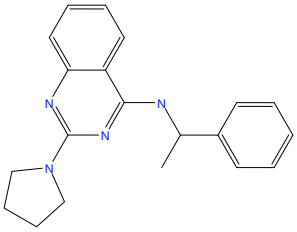
SSI-022
inhibition (%): 92.20



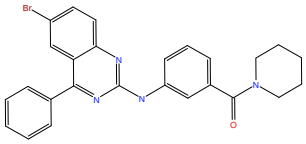
SSI-023
inhibition (%): 79.50



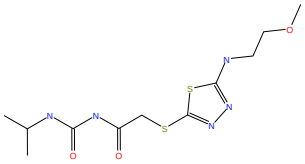
SSI-024
inhibition (%): 96.10



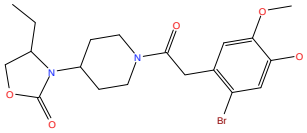
SSI-025
inhibition (%): 87.50



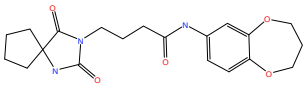
SSI-026
inhibition (%): 111.70



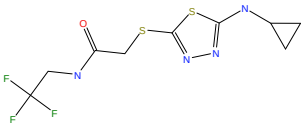
SSI-027
inhibition (%): 105.20



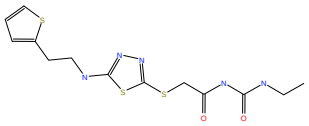
SSI-028
inhibition (%): 83.90



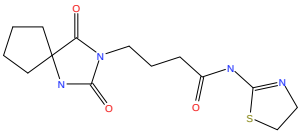
SSI-029
inhibition (%): 88.40



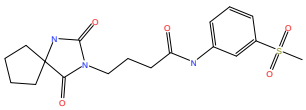
SSI-030
inhibition (%): 71.70



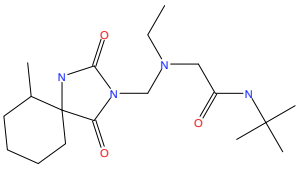
SSI-031
inhibition (%): 53.00



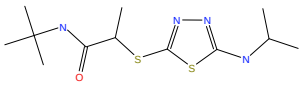
SSI-032
inhibition (%): 44.20



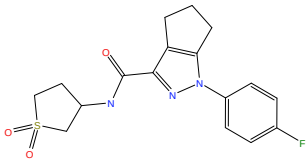
SSI-033
inhibition (%): 42.80



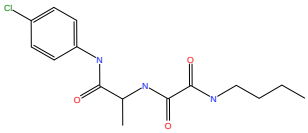
SSI-034
inhibition (%): 67.60



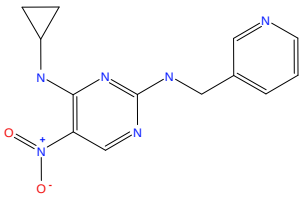
SSI-035
inhibition (%): 100.00



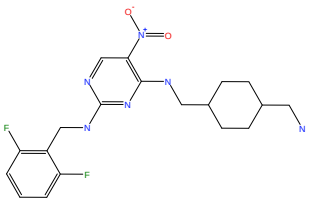
SSI-036
inhibition (%): 104.40



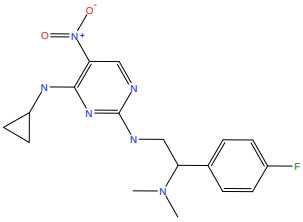
SSI-037
inhibition (%): 92.20



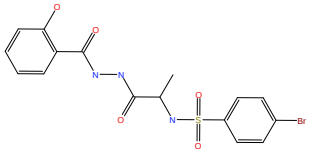
SSI-038
inhibition (%): 11.60



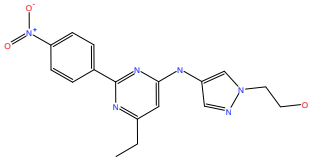
SSI-039
inhibition (%): 72.10



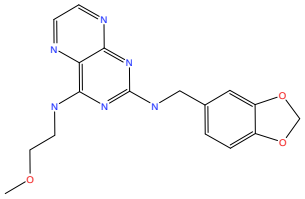
SSI-040
inhibition (%): 58.60



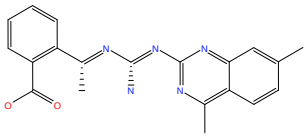
SSI-041
inhibition (%): 6.10



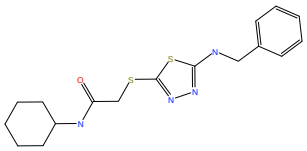
SSI-042
inhibition (%): 47.80



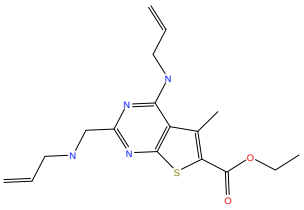
SSI-043
inhibition (%): 98.90



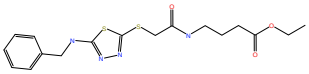
SSI-044
inhibition (%): 189.00



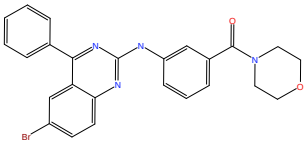
SSI-045
inhibition (%): 76.70



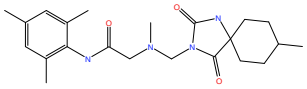
SSI-046
inhibition (%): 92.70



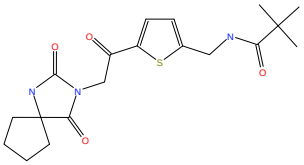
SSI-047
inhibition (%): 59.50



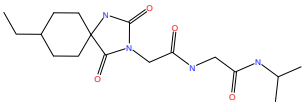
SSI-048
inhibition (%): 61.00



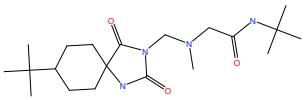
SSI-049
inhibition (%): 44.20



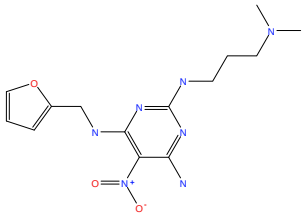
SSI-050
inhibition (%): 100.90



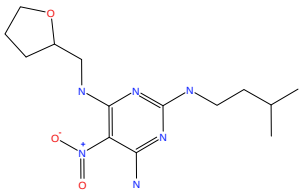
SSI-051
inhibition (%): 103.80



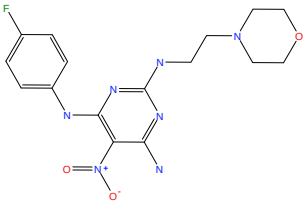
SSI-052
inhibition (%): 83.90



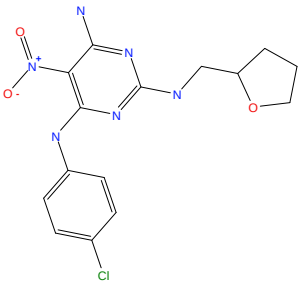
SSI-053
inhibition (%): 83.00



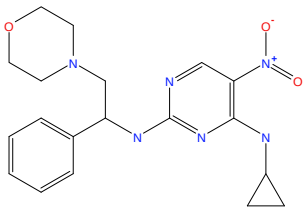
SSI-054
inhibition (%): 71.90



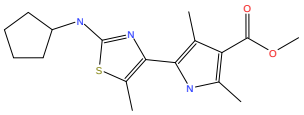
SSI-055
inhibition (%): 68.50



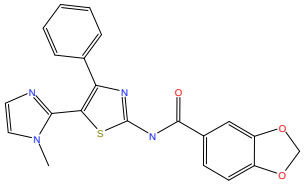
SSI-056
inhibition (%): 61.40



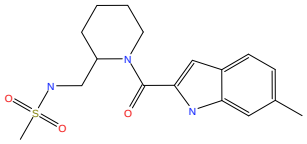
SSI-057
inhibition (%): 55.50



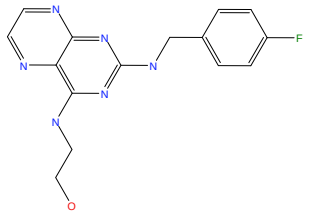
SSI-058
inhibition (%): 1.50



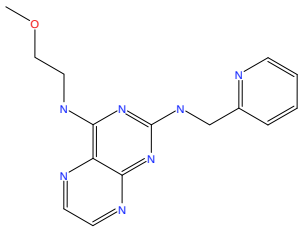
SSI-059
inhibition (%): 75.40



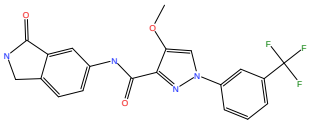
SSI-060
inhibition (%): 47.20



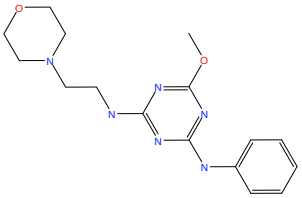
SSI-061
inhibition (%): 57.90



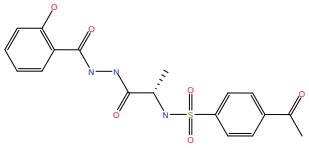
SSI-062
inhibition (%): 10.30



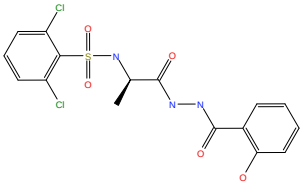
SSI-063
inhibition (%): 13.80



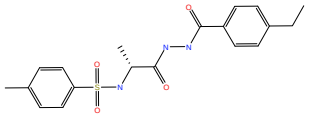
SSI-064
inhibition (%): 56.10



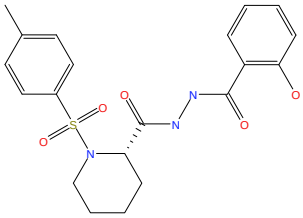
SSI-065
inhibition (%): 69.50



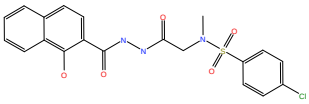
SSI-066
inhibition (%): 69.40



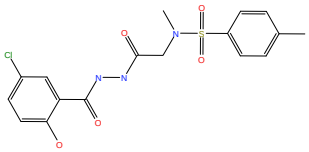
SSI-067
inhibition (%): 68.20



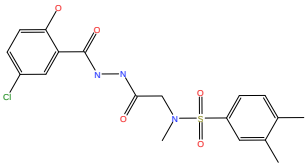
SSI-068
inhibition (%): 74.30



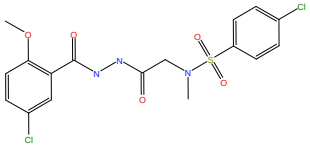
SSI-069
inhibition (%): 80.70



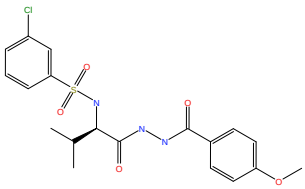
SSI-070
inhibition (%): 77.70



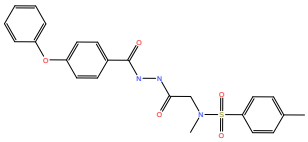
SSI-071
inhibition (%): 61.30



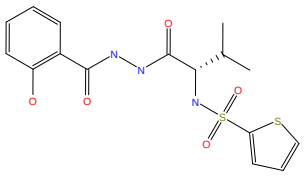
SSI-072
inhibition (%): 108.00



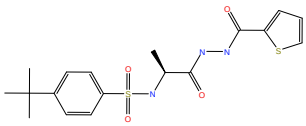
SSI-073
inhibition (%): 89.80



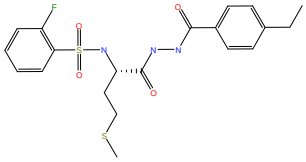
SSI-074
inhibition (%): 76.60



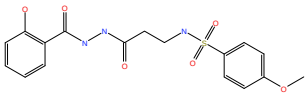
SSI-075
inhibition (%): 96.70



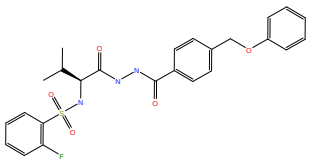
SSI-076
inhibition (%): 100.70



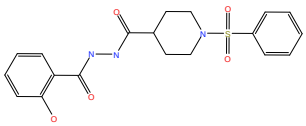
SSI-077
inhibition (%): 55.20



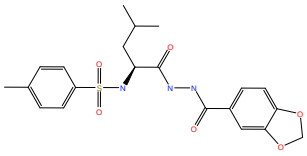
SSI-078
inhibition (%): 68.10



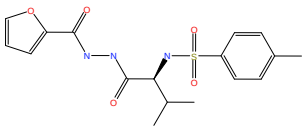
SSI-079
inhibition (%): 68.00



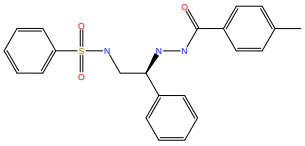
SSI-080
inhibition (%): 94.30



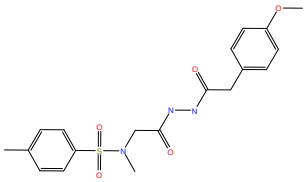
SSI-081
inhibition (%): 75.90



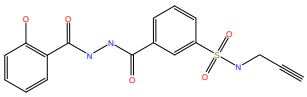
SSI-082
inhibition (%): 72.40



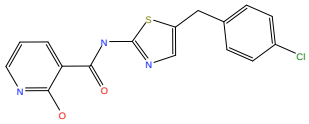
SSI-083
inhibition (%): 75.40



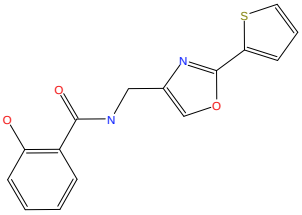
SSI-084
inhibition (%): 111.60



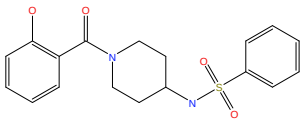
SSI-085
inhibition (%): 58.70



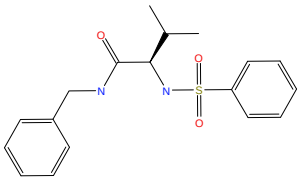
SSI-086
inhibition (%): 84.00



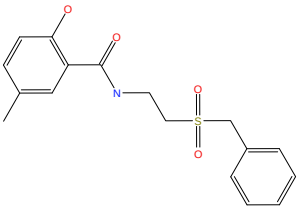
SSI-087
inhibition (%): 82.70



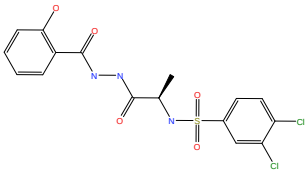
SSI-088
inhibition (%): 107.90



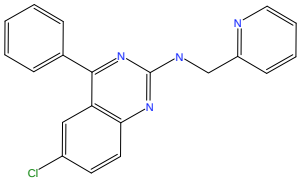
SSI-089
inhibition (%): 79.60



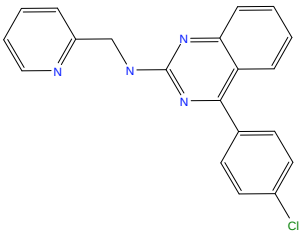
SSI-090
inhibition (%): 68.90



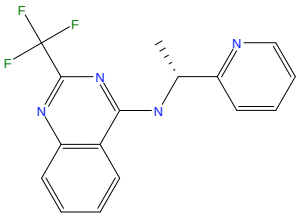
SSI-091
inhibition (%): 6.10



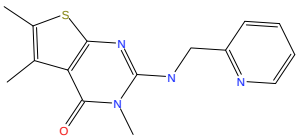
SSI-092
inhibition (%): 104.70



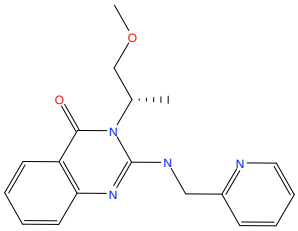
SSI-093
inhibition (%): 154.70



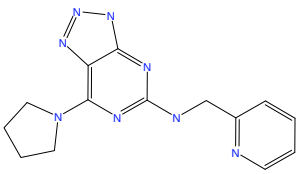
SSI-094
inhibition (%): 68.60



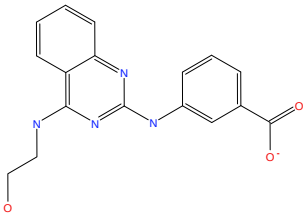
SSI-095
inhibition (%): 46.80



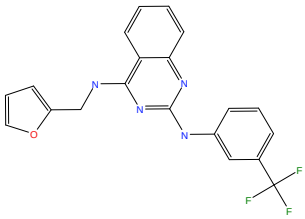
SSI-096
inhibition (%): 0.50



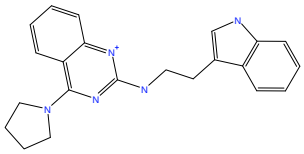
SSI-097
inhibition (%): 58.70



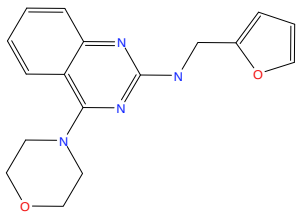
SSI-098
inhibition (%): 98.80



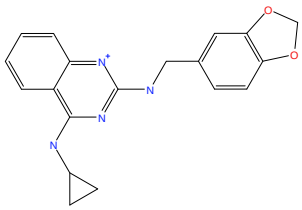
SSI-099
inhibition (%): 104.60



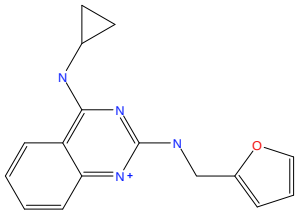
SSI-100
inhibition (%): 77.30



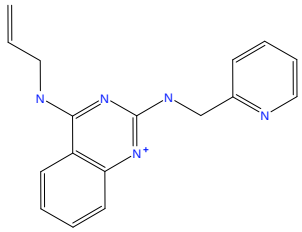
SSI-101
inhibition (%): 63.40



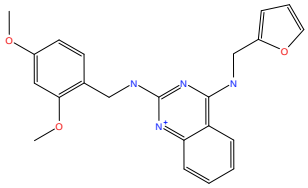
SSI-102
inhibition (%): 60.00



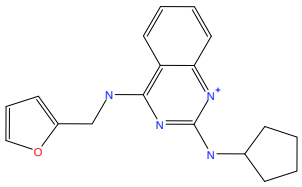
SSI-103
inhibition (%): 137.10



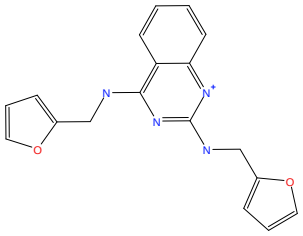
SSI-104
inhibition (%): 1.00



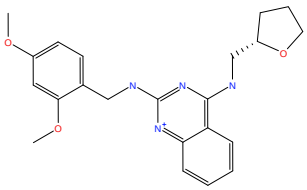
SSI-105
inhibition (%): 139.90



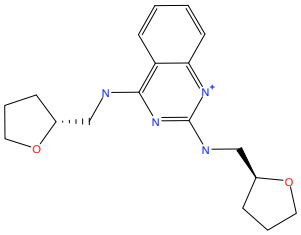
SSI-106
inhibition (%): 86.50



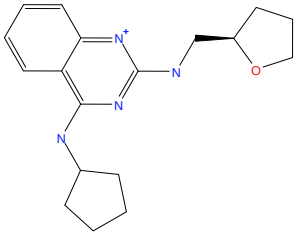
SSI-107
inhibition (%): 119.40



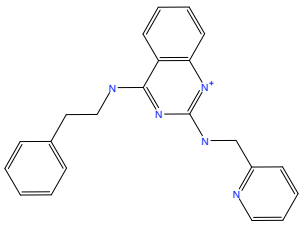
SSI-108
inhibition (%): 177.80



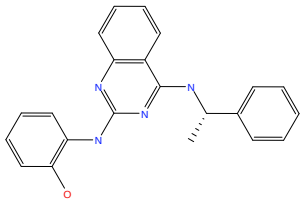
SSI-109
inhibition (%): 119.20



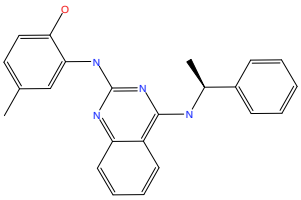
SSI-110
inhibition (%): 237.70



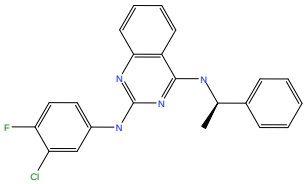
SSI-111
inhibition (%): 241.10



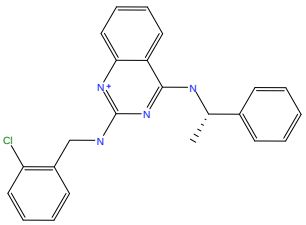
SSI-112
inhibition (%): 106.00



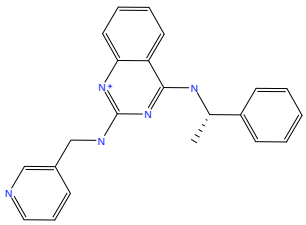
SSI-113
inhibition (%): 65.80



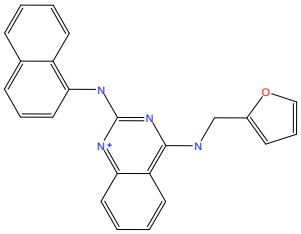
SSI-114
inhibition (%): 93.50



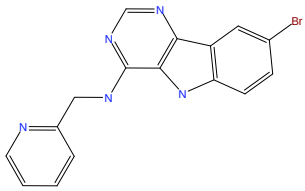
SSI-115
inhibition (%): 139.40



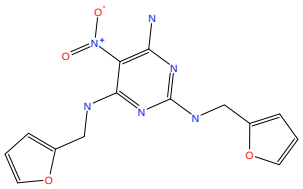
SSI-116
inhibition (%): 210.80



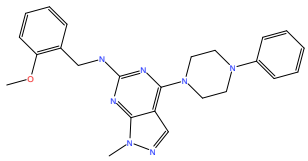
SSI-117
inhibition (%): 131.20



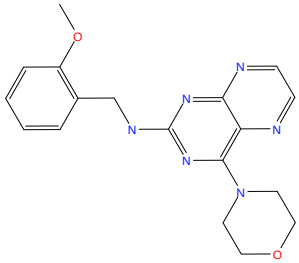
SSI-118
inhibition (%): 112.90



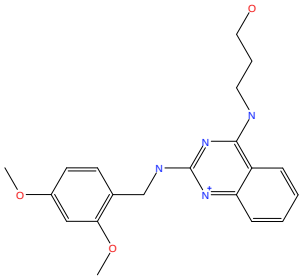
SSI-119
inhibition (%): 92.10



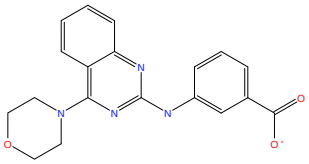
SSI-120
inhibition (%): 69.20



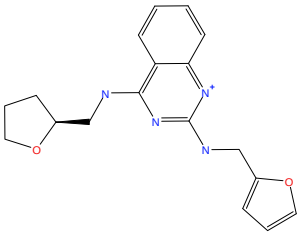
SSI-121
inhibition (%): 159.10



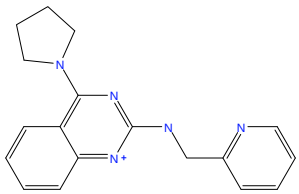
SSI-122
inhibition (%): 151.00



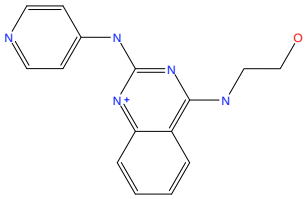
SSI-123
inhibition (%): 102.40



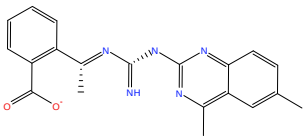
SSI-124
inhibition (%): 99.00



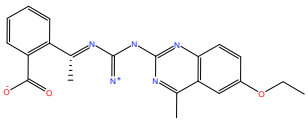
SSI-125
inhibition (%): 121.80



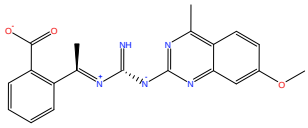
SSI-126
inhibition (%): 75.00



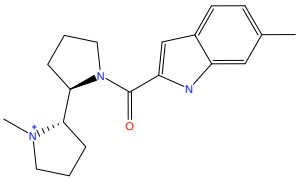
SSI-127
inhibition (%): 484.20



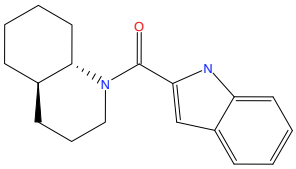
SSI-128
inhibition (%): 547.10



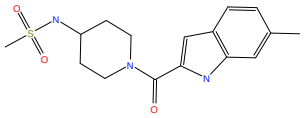
SSI-129
inhibition (%): 391.30



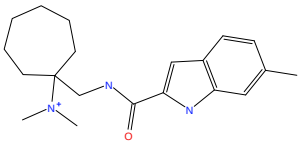
SSI-130
inhibition (%): 185.00



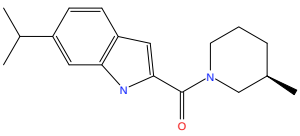
SSI-131
inhibition (%): 149.00



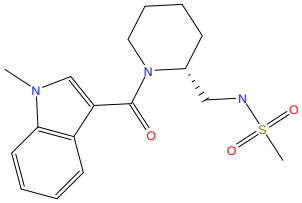
SSI-132
inhibition (%): 190.40



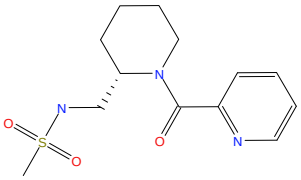
SSI-133
inhibition (%): 338.00



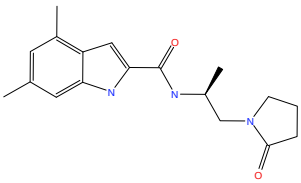
SSI-134
inhibition (%): 223.70



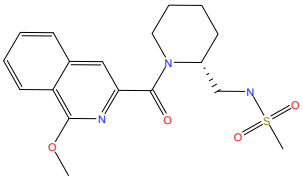
SSI-135
inhibition (%): 153.30



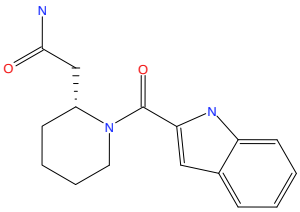
SSI-136
inhibition (%): 165.00



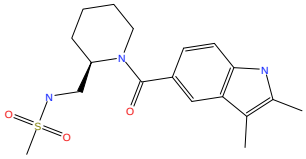
SSI-137
inhibition (%): 235.60



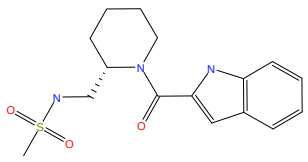
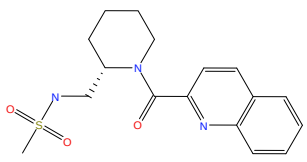
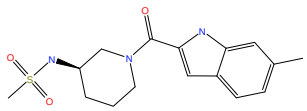
SSI-138
inhibition (%): 254.80



SSI-139
inhibition (%): 76.90



SSI-140
inhibition (%): 98.60

 SSI-141 inhibition (%): 180.70	 SSI-142 inhibition (%): 141.20	 SSI-143 inhibition (%): 125.30

References:

1. F. K. Pettit, E. Bare, A. Tsai and J. U. Bowie, *Journal of molecular biology*, 2007, **369**, 863-879.
2. N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon, *Proceedings of the National Academy of Sciences of the United States of America*, 2001, **98**, 10037-10041.
3. P. C. Hawkins and A. Nicholls, *Journal of chemical information and modeling*, 2012, **52**, 2919-2936.
4. P. C. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson and M. T. Stahl, *Journal of chemical information and modeling*, 2010, **50**, 572-584.
5. B. O. Villoutreix, N. Renault, D. Lagorce, O. Sperandio, M. Montes and M. A. Miteva, *Current protein & peptide science*, 2007, **8**, 381-411.
6. G. Jones, P. Willett and R. C. Glen, *Journal of computer-aided molecular design*, 1995, **9**, 532-549.
7. H. Gohlke, M. Hendlich and G. Klebe, *Journal of molecular biology*, 2000, **295**, 337-356.
8. G. Neudert and G. Klebe, *Journal of chemical information and modeling*, 2011, **51**, 2731-2745.
9. A. Voet, F. Berenger and K. Y. J. Zhang, *PloS one*, 2013, **8**, e75762.
10. Y. Uchimura, M. Nakao and H. Saitoh, *FEBS letters*, 2004, **564**, 85-90.
11. N. Rouleau, J. Wang, L. Karras, E. Andrews, M. Bielefeld-Sevigny and Y. Chen, *Analytical biochemistry*, 2008, **375**, 364-366.
12. T. Kigawa, T. Yabuki and S. Yokoyama, *Tanpakushitsu kakusan koso. Protein, nucleic acid, enzyme*, 1999, **44**, 598-605.
13. T. Matsuda, S. Koshiba, N. Tochio, E. Seki, N. Iwasaki, T. Yabuki, M. Inoue, S. Yokoyama and T. Kigawa, *Journal of biomolecular NMR*, 2007, **37**, 225-229.
14. J. Yokoyama, T. Matsuda, S. Koshiba, N. Tochio and T. Kigawa, *Analytical biochemistry*, 2011, **411**, 223-229.
15. F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer and A. Bax, *Journal of biomolecular NMR*, 1995, **6**, 277-293.
16. N. Kobayashi, J. Iwahara, S. Koshiba, T. Tomizawa, N. Tochio, P. Guntert, T. Kigawa and S. Yokoyama, *Journal of biomolecular NMR*, 2007, **39**, 31-52.
17. B. A. Johnson, *Methods Mol Biol*, 2004, **278**, 313-352.