

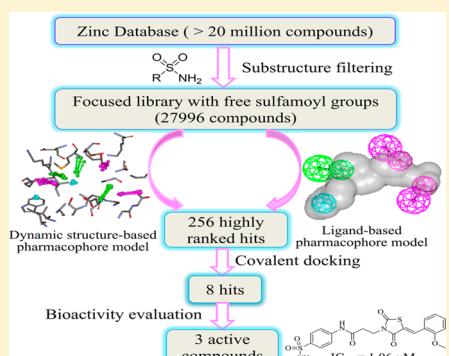
Effective Virtual Screening Strategy toward Covalent Ligands: Identification of Novel NEDD8-Activating Enzyme Inhibitors

Shengping Zhang,^{†,‡,⊥} Jiani Tan,^{†,§,⊥} Zhonghui Lai,^{†,‡} Ying Li,^{||} Junxia Pang,^{†,‡} Jianhu Xiao,^{†,‡} Zhangjian Huang,^{†,‡} Yihua Zhang,^{†,‡} Hui Ji,^{*,†,§} and Yisheng Lai^{*,†,‡}

[†]State Key Laboratory of Natural Medicines, [‡]Center of Drug Discovery, [§]Department of Pharmacology, and ^{||}School of Pharmacy, China Pharmaceutical University, Nanjing 210009, Jiangsu, China

Supporting Information

ABSTRACT: The NEDD8-activating enzyme (NAE) is an emerging target for cancer therapy, which regulates the degradation and turnover of a variety of cancer-related proteins by activating the cullin-RING E3 ubiquitin ligases. Among a limited number of known NAE inhibitors, the covalent inhibitors have demonstrated the most potent efficacy through their covalently linked adducts with NEDD8. Inspired by this unique mechanism, in this study, a novel combined strategy of virtual screening (VS) was adopted with the aim to identify diverse covalent inhibitors of NAE. To be specific, a docking-enabled pharmacophore model was first built from the possible active conformations of chosen covalent inhibitors. Meanwhile, a dynamic structure-based pharmacophore was also established based on the snapshots derived from molecular dynamic simulation. Subsequent screening of a focused ZINC database using these pharmacophore models combined with covalent docking discovered three novel active compounds. Among them, compound LZ3 exhibited the most potent NAE inhibitory activity with an IC_{50} value of $1.06 \pm 0.18 \mu\text{M}$. Furthermore, a cell-based washout experiment proved the proposed covalent binding mechanism for compound LZ3, which confirmed the successful application of our combined VS strategy, indicating it may provide a viable solution to systematically discover novel covalent ligands.



1. INTRODUCTION

Targeting regulators of cellular protein homeostasis has become an emerging therapeutic field for a broad array of diseases, including neurodegenerative disorders, viral infection, and cancer.^{1–3} The ubiquitin-proteasome system (UPS), comprising the ubiquitin and ubiquitin-like protein (UBL) pathways, is the primary mediator of protein degradation and controls the turnover of a large range of proteins in eukaryotic cells.⁴ The first drug targeting the UPS is bortezomib which was approved by FDA in 2003 for the treatment of multiple myeloma and mantle cell lymphoma.^{5,6} The case of bortezomib suggested the diverse set of enzymes in the UPS could provide a wide platform for creating novel cancer therapeutics.

Among nine classes of UBLs identified, the neural precursor cell-expressed developmentally down-regulated 8 (NEDD8) plays an essential role in cell progression and signal transduction. Like the ubiquitin pathway, NEDD8 is initially activated by NEDD8 activating enzyme (NAE) in an ATP-dependent reaction and then transferred to Ubc12, the NEDD8 conjugating enzyme. Finally, with the coordinated action of Ubc12 and SCCRO (DCN1), the scaffold-type E3 ligase for cullin NEDDylation, NEDD8 is attached to a conserved lysine near the C-terminal end of the cullin protein.⁷ The NEDDylation of cullin proteins is of great significance to activate cullin-RING ligases (CRLs), a subclass of ubiquitin E3 ligases which controls degradation of a specific subset of

proteins, such as cancer-related proteins Cdt1 and IκBα.⁸ Additionally, a recent study has also demonstrated there is a remarkable increase in the level of NEDDylation observed in malignant cancer cells.⁹ Thus targeting the NEDD8 pathway, particularly NAE, for cancer therapy has raised deep interest world wide.^{10–13}

MLN4924 is the first and only inhibitor of NAE that enters clinical trials at present for the treatment of several types of cancer, especially hematologic malignancies. It was identified by the Millennium Company as a result of high throughput screening and relentless optimization efforts.¹⁴ MLN4924 is a potent and selective NAE inhibitor with an AMP mimetic scaffold, which is the natural substrate of NAE. In light of the high structural similarity between MLN4924 and AMP, it is quite reasonable to assume they have a similar binding pattern and mechanism when interacting with NAE. Then further extensive studies confirmed this postulation by revealing the key for inhibition of NAE resides in the formation of a NEDD8-MLN4924 covalent adduct, which resembles the NEDD8-AMP. The adduct effectively inhibits the NAE by occupying its ATP binding sites, obstructing NEDD8 trans-thiolation to E2, and finally leading to inactivation of the entire NEDD8 pathway.¹⁵ However, despite its promising therapeutic

Received: April 2, 2014

potential, treatment-emergent drug resistance caused by heterozygous mutations in the ATP binding pocket of NAE has emerged in cell and xenograft models of cancer,¹⁶ which presses the urgent need to develop novel NAE inhibitors.

Virtual screening (VS) technologies, such as pharmacophore modeling and molecular docking simulation, have earned great reputation in rapidly identifying potential lead compounds at relatively lower cost of economic investment.^{17–20} In the case of NAE, considering the significance of covalent binding and the inefficiency of known VS strategy for covalent ligands, a new hierarchical VS method which integrated covalent linkage throughout VS process was proposed to identify potential covalent NAE inhibitors. Specifically, two pharmacophore models derived from the information of known covalent inhibitors and the molecular dynamic (MD) simulation of the NAE-NEDD8 complex were generated and validated by widely accepted methods, which were then used as 3D queries to search a focused library for potent leads. Hit compounds were further assessed for their binding rationality with NAE via covalent docking, and eight compounds were finally selected for experiment verification. Gratifyingly three of the eight hits exhibited potent inhibitory activity against NAE in a cell-free system, and the covalent binding mechanism was confirmed by further washout experiment for the most active compound LZ3. The results support the successful application of our proposed VS strategy in the identification of covalent inhibitors.

2. MATERIALS AND METHODS

2.1. Docking Enabled Common Feature Pharmacophore Generation. Seven known covalent inhibitors of NAE were selected as the training set to establish ligand-based pharmacophore (LBP) based on their NAE inhibitory activities and structural diversity.^{13,21–23} All the molecules were sketched in ChemDraw Ultra 12.0 and converted into 3D structures when exported into Discovery Studio 2.5 (DS). Then compounds were subjected to local minimization using the CHARMM force field.²⁴ The resulting structures were then docked into the NAE receptor (Protein database code: 3GZN) by Gold 5.0 to identify their possible bioactive conformations (Docking procedures will be described in the following section.). The postdocking conformations for active compounds were utilized to establish the common feature pharmacophore using Hiphop²⁵ in DS. Accordingly, in this building process the parameter for conformation sampling was turned off. The Principal and MaxOmitFeat values were set to 2 and 0 for MLN4924, 1 and 1 for all the other remaining molecules in the training set. Hydrogen bond donor (HBD), hydrogen bond acceptor (HBA), hydrophobic, and aromatic ring (AR) features were considered as important chemical features of NAE inhibitors based on the Feature Mapping result. The Minimum Features were set to 4 to direct the production of valid pharmacophores. Ten hypotheses were generated, and the best hypothesis was selected based on the ligand–receptor interactions as well as decoy test. A shape constraint was then added based on the volume of the template molecule MLN4924.

2.2. Structure-Based Pharmacophore Generation. The X-ray crystallographic structure of NAE (PDB code: 3GZN) was downloaded from the PDB. The NAE regulatory subunit was removed from the structure, because it is far away from the active site of MLN4924. The resulting complex was then prepared by using the Prepare Protein utility in DS with missing atoms fixed and hydrogen atoms added in a

physiological condition. Before the MD, the covalent bond between Gly76 of NEDD8 and MLN4924 was broken. The missing oxygen atom was first added based on the available moiety and then subjected to sequent local minimization to ensure a proper start conformation for MD simulation.

The prepared structure was immersed in a TIP3P water box which extended 10 Å from the surface of the enzyme. Sodium and chloride ions were added to neutralize the system and to mimic physiological conditions. Then 5000 steps of steepest descent (SD) minimization combined with 5000 steps of conjugate gradient (CG) minimization were first performed for only waters and counterions and then for the whole system with the CHARMM force field. The relaxed system was subsequently heated to 300 K for 50 ps with 5 kcal mol⁻¹Å⁻² harmonic position restraints on the heavy atoms then followed by 300 ps unconstrained equilibrium in NPT ensemble. Finally the MD production run was performed for another 6 ns. During the MD, the SHAKE algorithm was employed to constrain all bonds involving hydrogen atoms with a 2 fs time step.²⁶ The nonbonded interaction cutoff was set to 12 Å, and the particle mesh Ewald method (PME) was applied to compute long-range electrostatic interaction.²⁷ The periodic boundary conditions with NPT ensemble were adopted throughout the equilibration and production phase. All simulations involved in the MD were carried out with the CHARMM²⁸ suite available in DS. A total of 100 snapshots were collected from the MD trajectory with an interval of 40 ps after the system was stable. Then the 100 snapshots were clustered to 10 bins from which an average conformation would be selected to represent each cluster based on the root-mean-square deviation (RMSD) values of backbone atoms. Finally the representative 10 snapshots were aligned to the structure in the last snapshot of the entire 6 ns MD trajectory with the aim to achieve a common frame of reference.

After removing water, ions, and ligand, each snapshot was used to derive a structure-based pharmacophore (SBP) using the Ludi algorithm²⁹ in DS. This protocol first creates an interaction map inside a receptor active site sphere which consists of three types of features: HBA, HBD, and hydrophobic feature. Then a pharmacophore was generated from the map after hierarchical clustering of overlapping features. A total of 10 pharmacophore models from all snapshots were eventually overlaid, and the features that were conserved throughout the MD trajectory (present in at least 5 models) were retained. The center of each feature cluster was determined by the average position of its contributing elements, and the radius was given by the RMSD of contributing elements with a multiplication factor of 2.5. In this routine, protein flexibility was implicitly accommodated by loosening the requirement for features in the dynamic region so that the resulting pharmacophore model would focus on the real conserved interaction feature, thus increasing the diversity of identified ligands. Besides, excluded volumes were added to account for the steric hindrance according to the position of protein atoms within 3 Å of the ligand in each snapshot. The final position of each excluded volume in the consensus pharmacophore was defined by the average position of each atom, and a radius of 1.5 Å was given to all atom types.

2.3. Validation of Pharmacophore Models. Pharmacophore evaluation is critical to assess the capability of generated hypotheses to distinguish active compounds from inactives. Here the Ligand Pharmacophore Mapping and Screen Library modules in DS were utilized in the validation process to screen

a decoy database which contains 30 known active NAE inhibitors and 1000 decoys. The reliability of pharmacophore models was measured by two different matrices: enrichment factor (EF) and goodness of hit (GH).³⁰ The EF value was calculated as fraction of retrieved actives in the hit lists divided by the fraction of total actives in the database, and the GH value was computed through the following equation

$$EF = (Ha \times D) / (Ht \times A) \quad (1)$$

$$GH = [(Ha \times (3A + Ht)) / (4Ht \times A)] \times [1 - (Ht - Ha) / (D - A)] \quad (2)$$

where Ha and Ht represent the number of positives in the hits and the total number of hits, respectively, and A and D refer to the number of actives and the total number of molecules in the database correspondingly.

The 30 active compounds in the decoy database were collected from published literature and patents.^{10–13,21–23} The decoy databases were built by following the rules similar to the strategy Shoichet and Irwin³¹ adopted in construction of the decoy sets from ZINC.³² Namely, the decoys should have similar physical properties with seeded known NAE inhibitors and can also be distinguished topologically. With this concept in mind, three steps were employed to choose the proper decoy molecules. First, compounds' similarity with 30 annotated NAE inhibitors were scaled by the Tanimoto coefficient³³ using the extended-connectivity fingerprints³⁴ (FCFP_6). Compounds with the Tanimoto coefficient less than 0.6 were selected. Second, five physical properties including molecular weight, number of HBAs and HBDs, number of rotatable bonds, and AlogP were calculated for both known NAE inhibitors and molecules obtained from last steps. The choices were made only when the calculated physical properties of certain molecules were close to those of any annotated NAE inhibitors. Besides, in order to mimic the unbalanced nature of inactives and actives in the real database screening, a high ratio of decoy molecules versus true actives (100:3) was adopted in this research which led to a total of 1000 decoys from ZINC. Finally, the identified 1000 decoy molecules were combined with 30 annotated NAE inhibitors to form the decoy database. The ligand flexibility of all 1030 compounds was incorporated in the modeling by an exhaustive sampling of different conformations representing the energy accessible space to the molecule. This work was completed by using the Diverse Conformation Generation protocol in DS with the Best option. The energy threshold was set to 20 kcal/mol, and the Maximum Conformations was set as 255.

2.4. Covalent Docking. Covalent docking was performed under the covalent docking mode available in Gold 5.0. In covalent mode, the program assumed that there is just one atom linking the ligand to the protein. Both protein and ligand files were set up with the link atom included. In the docking routine, the link atom in the ligand was forced to fit onto the link atom in the protein. Finally an angle-bending energy term for the link atom was involved in the calculation of the fitness score to ensure the geometry of the bound ligand was correct.³⁵ All the ligands containing sulfamoyl groups were subjected to the aforementioned covalent docking in Gold with the modified version of Goldscore and Aspscore. Both protein and ligands were prepared with the link atom, namely the nitrogen atoms of sulfamoyl substructure in ligands, according to Gold user manual. All the other parameters involved in covalent docking

remained default. The reliability of the chosen parameters was subsequently validated by redocking and another decoy database which contains 30 covalent NAE inhibitors and 1000 noninhibitors with free sulfamoyl groups. The decoy molecules were selected from ZINC with the same strategy used above.

2.5. Focused Library Design and Screening. The publicly available database, ZINC, which contains 22,724,825 commercially purchasable compounds, was screened using our validated pharmacophore models. Since the covalent linkage between ligands and NAE were defined only between the sulfamoyl group in the ligand and the carbonyl group of Gly76, the ZINC database was first filtered using the Align to Selected Substructure protocol in DS to obtain all the compounds with free sulfamoyl groups. The focused library which contains a total of 27996 molecules was then built of multiconformation by using the 'Build 3D Database' utility in DS (best option, the maximum number of conformers = 255) and served as the primary source for subsequent screening.

2.5.1. Pharmacophore Screening. The Search 3D Database module with the fast search method was employed when screening with LBP, and for SBP searching, we adopted the Screen library module using the flexible searching method with five of seven features required. The resulting hits were ranked according to their geometry fitting with the two hypotheses which was indicated as the Fit value. Then compounds with the high Fit value were filtered by Lipinski's rule of five and Veber's rule to eliminate hits with undesirable physiological properties. Ultimately only molecules with a Fit value above 2.5 in LBP searching or greater than 2 in SBP screening were selected and retained for molecular docking study.

2.5.2. Docking Screening. All the molecules passing the pharmacophore screening were combined and subjected to covalent screening to further narrow down the hit list. Gold 5.0 with fine-tuned parameters mentioned above were applied in this process and Goldscore and ASP were adopted to rank the docking results. Compounds with more than one free sulfamoyl group were separately docked with all possible linking groups, and the ones with the highest docking score were selected for further consideration. Finally the top-ranked hits by Goldscore and ASP were extracted and visually inspected for hydrogen-bonding interaction with important residues. Molecules with high scores and favorable hydrogen bonds with essential residues were picked out for biological tests.

2.6. Bioassay. **2.6.1. Materials and Antibodies.** All the hit compounds were purchased from Enamine, Life Chemicals, and Chembridge and dissolved in DMSO to a concentration of 100 mM. The final concentration of DMSO was less than 0.1% (v/v) in all experiments. 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) was purchased from Sigma Chemical Company (St. Louis, USA). The NEDD8 conjugation initiation kit and the primary antibody against Ubc12 were purchased from Boston Biochem (Cambridge, MA, USA). Primary antibodies against β -actin and the horseradish peroxidase-conjugated antirabbit antibody were supplied by Bioworld Technology Inc. (MN, USA).

2.6.2. Cell Lines. MCF-7, Caco-2, and Bcl-7402 cells were obtained from the Shanghai Institute of Cell Biology (Shanghai, China). All cells were maintained in HG-Dulbecco's modified Eagle's medium (Gibco, Invitrogen). The medium for the culture of cell lines was supplemented with 10% fetal calf serum (PAA, Australia), 100 IU/mL penicillin, and 100 IU/mL streptomycin (Amresco, USA). Both cell lines were cultured at

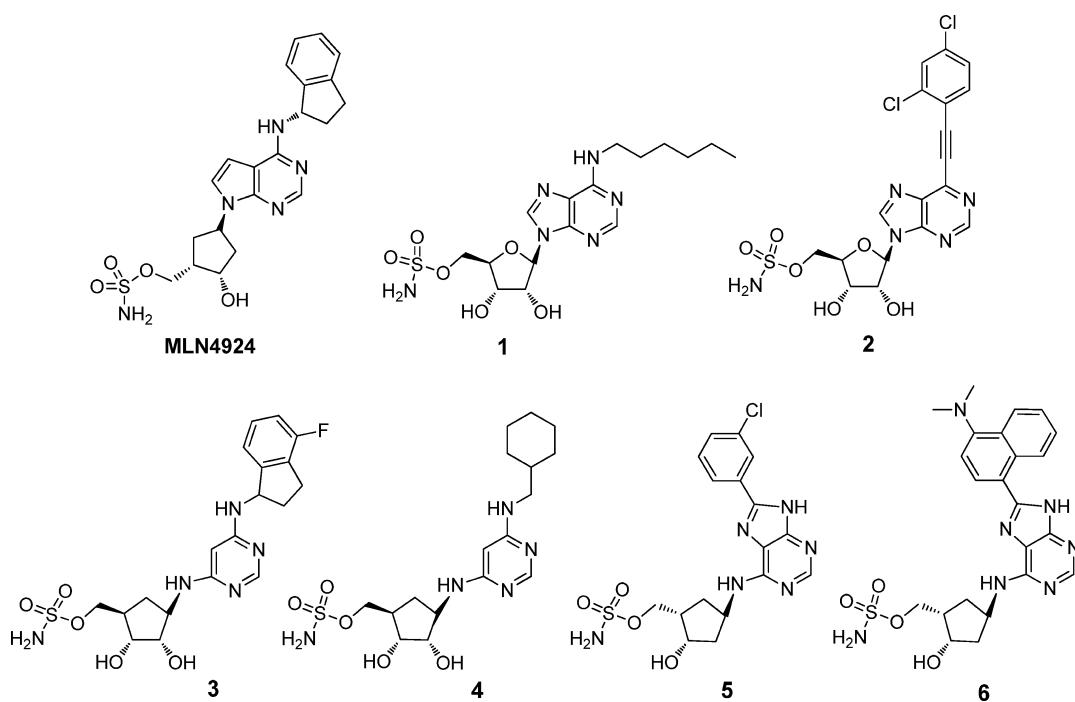


Figure 1. Chemical structure of the training set compounds.

37 °C in a humidified atmosphere consisting of 5% CO₂ and 95% air.

2.6.3. Cell-Free NAE Activity Assay. E1-like NAE activity assay was measured using the NEDD8 conjugation initiation kit according to the manufacturer's instructions. Briefly, NAE, NEDD8, Ubc12, and indicated concentrations of tested compounds were mixed in the reaction buffer and incubated for 10 min, and then Mg-ATP solution was added to initiate the reaction. Specifically, for hit compounds the concentration gradients were set as 0, 0.938, 1.875, 3.75, 7.5, 15, and 30 μM, while the concentration gradients for MLN4924 were 0, 3.125, 6.25, 12.5, 25, 50, and 100 nM. After having been incubated at 37 °C for 2 h, the reaction was terminated with EDTA. The Ubc12 and Ubc12-NEDD8 uncoupling protein levels were determined by Western blot analysis using the primary antibody against Ubc12.

2.6.4. Cytotoxicity Assay. The cytotoxic activity of compounds *in vitro* was determined using the MTT assay. Cells were plated in a 96-well plate at a density of 5×10^3 cells per well. After incubation overnight, cells were then treated with various concentrations of tested compounds for the indicated time. Twenty μL of MTT (5 mg/mL in PBS) was added to the wells and incubated for 4 h. The formed formazan crystals were dissolved in 150 μL of DMSO and measured using a microplate spectrophotometer (Tecan, Switzerland) at the test wavelength of 570 nm. The inhibitory rate was calculated using the following equation

$$\text{I\%} = (1 - \text{A}_{\text{exp}} / \text{A}_{\text{con}}) \times 100\% \quad (3)$$

where A_{exp} and A_{con} denote the absorbance of the test substances and solvent control, respectively.

To study the cell-based NAE activity of compounds, cells were pretreated with indicated concentrations of compounds for 3 days, and were lysed in RIPA buffer. The concentration was measured by BCA protein assay (Beyotime, Nanjing, China). The Ubc12 and Ubc12-NEDD8 uncoupling protein levels were determined by Western blot analysis.

2.6.5. Western Blot Analysis. The proteins in the cell lysates (or the reaction mixture) were separated on a SDS-PAGE and transferred to a polyvinylidene difluoride membrane (Millipore, MA, USA). The membranes were blocked with 5% nonfat milk in TBST buffer (20 mM Tris (pH 8.0), 150 mM NaCl, and 0.1% Tween-20) and incubated overnight at 4 °C with the Ubc12 primary antibodies (The primary antibody against Ubc12 was used to pick up the level of both Ubc12 and Ubc12-NEDD8 thioester product.) followed by incubation with the horseradish peroxidase-conjugated secondary antirabbit antibody for 2 h. The proteins were visualized using an ECL KeyGEN system (KeyGEN Biotechnology, Nanjing, China) and scanned with a Clinix ChemiScope chemiluminescence imaging system (Gel Catcher 2850, China). The relative optical densities of specific proteins were determined using a ChemiScope analysis program.

3. RESULTS AND DISCUSSION

3.1. Ligand-Based Pharmacophore Modeling. Common Feature Pharmacophores were generated by the HipHop algorithm available in DS, using 7 known NAE inhibitors as the training set (Figure 1). Since the HipHop algorithm is designed to identify the common chemical features of active molecules based on their bioactive conformations superimposition, proper starting conformation for each training set compound is essential for a successful pharmacophore modeling. Normally, an extensive sampling of conformation space is required for each building molecule to discover their potential active conformations before the pharmacophore generation. In this case, several possible conformation overlays could be identified during model generation stage where multiple pharmacophores with distinct composition and arrangement could be generated. Thus a wise decision is required to choose the optimal one. Here, we adopted a different strategy by incorporating covalent docking to address the active conformation issue. Seven compounds in the training set were covalently docked to the

structure of NAE using Gold, and for each compound the top pose ranked by Goldscore was selected for subsequent pharmacophore modeling. The reliability of docking experiment was examined by redocking the native ligand-MLN4924 back to the NAE structure. The RMSD value between the top-ranked pose and the crystallographical conformation was 0.78 Å, suggesting a good performance in conformation prediction.

In the modeling of common feature pharmacophore, the Principal and MaxOmitFeat values define the number of molecules which must completely or partially map to the hypothesis. In our case, all seven training set compounds are all highly potent covalent inhibitors of NAE with $IC_{50} < 100$ nM. Among them, MLN4924 which entered phase I clinic trials has been well established as the reference in NAE inhibition. Thus MLN4924 was assigned a Principal value of 2 to instruct HipHop to consider all the pharmacophore features it has in pharmacophore building, and a Principal value of 1 was given to the rest of the compounds in the training set to ensure they will be at least once mapped by each generated hypothesis. On the other hand, the MaxOmitFeat value of 0 was assigned to MLN4924 which forces mapping of all features of this compound, while others were allowed to miss one feature in any produced pharmacophore by giving them the MaxOmitFeat value of 1.

Ultimately, 10 qualitative pharmacophores were constructed, and for each created hypothesis, all compounds in the training set were well fitted with no partial mapping or missing features. As one can see in Table 1, the generated pharmacophore models can be clustered to three categories based on their feature composition and space organization. Category I included the top-ranked two hypotheses which consisted of two HBDs, one HBA, and one Hydrophobic feature. Category II, however, contained only one hypothesis which was comprised of two HBAs, one HBD, and one Hydrophobic

feature. The last seven hypotheses with two HBDs, one HBA, and one AR were clustered to form Category III. Since within each category, all the hypotheses were ranked with identical score and pharmacophore features in each hypothesis were also spatially arranged in a similar pattern, we decided to automatically choose the first emerging model in each category as representatives for further analysis. Finally, the three remaining pharmacophore models, Hypo1, Hypo3 and Hypo4, were mapped to the crystal structure of MLN4924 to examine whether they were a true reflection of ligand-protein interactions. The mapping results were shown in Figure 2. The major difference of Hypo1 and Hypo3 resided in the mapping of the terminal amino group of MLN4924 which would serve as a warhead in the covalent binding event. In Hypo1, the terminal amino group of MLN4924 was successfully mapped to a HBD, while in Hypo3 this critical feature failed to be expressed in any forms. On the other hand, Hypo1 and Hypo4 differentiated from each other only by one feature (HBA-AR). It seemed that the HBA feature captured the hydrogen bonding interaction between MLN4924 and Ile148, while the R feature failed to express any interaction. Therefore Hypo1 was selected as the best common feature pharmacophore model that will be subjected to the following evaluation.

3.2. Dynamic Structure-Based Pharmacophore Modeling. In order to derive information from the NAE structure, a SBP model was constructed from the crystallographical structure of NAE complexed with NEDD8 and MLN4924 (PDB code: 3GZN). Since MLN4924 is covalently bound to the NAE-NEDD8 complex, covalent binding is usually considered to require an initial association, in which the ligand at first noncovalently interacts with the protein in a pose that facilitates the reaction.³⁶ Thus this binding is largely controlled by the access of the ligand to the binding site in a suitable pose. Based on this theory, a MD study was performed with the aim to investigate the variation of noncovalent interaction between ligand and receptor. In addition, protein flexibility is another issue that exerts heavy influence on structure-based drug design, and it is always insufficient to cover all energetically accessible chemical space of protein from a single, rigid representation.³⁷ Carlson et al. has reported a multiple protein structure (MPS) methodology through which receptor flexibility can be integrated into the pharmacophore modeling with a trajectory of MD conformations portraying a more comprehensive presentation of protein.^{38–40} Therefore, a similar strategy was employed in our study of NAE, and an ensemble of representative conformations from the MD running was collected as inputs for pharmacophore generation.

The 6 ns MD simulation results were indicated in Figure 3, the system, especially the binding site area (residues within 15 Å of ligand), was not stable until 2 ns. Thus snapshots only after 2 ns were used for further analysis. A total of 100 conformations originating from the last 4 ns were extracted from the MD trajectory and then clustered to 10 bins based on their backbone atoms RMSD values. Finally, ten representative conformations which were closest to the average structure of each bin were selected and overlaid to achieve the same frame of reference.

An interaction map inside the active site of NAE comprising HBA, HBD, and hydrophobic features was created using Ludi for each complex, and ten SBP models were subsequently generated from the Ludi interaction map by a hierarchical clustering methodology. The resulting 10 pharmacophore models were then superimposed together to identify features

Table 1. Details of the Top Ten Hypotheses Generated by HipHop

hypothesis name	features ^a	rank ^b	direct hit ^c	partial hit ^c	max. fit
Hypo1	H, HBD, HBD, HBA	82.964	1111111	0000000	4
Hypo2	H, HBD, HBD, HBA	82.964	1111111	0000000	4
Hypo3	H, HBD, HBA, HBA	82.801	1111111	0000000	4
Hypo4	R, H, HBD, HBD	82.524	1111111	0000000	4
Hypo5	R, H, HBD, HBD	82.524	1111111	0000000	4
Hypo6	R, H, HBD, HBD	82.524	1111111	0000000	4
Hypo7	R, H, HBD, HBD	82.524	1111111	0000000	4
Hypo8	R, H, HBD, HBD	82.524	1111111	0000000	4
Hypo9	R, H, HBD, HBD	82.524	1111111	0000000	4
Hypo10	R, H, HBD, HBD	82.524	1111111	0000000	4

^aH: hydrophobic; HBD: hydrogen bond donor; HBA: hydrogen bond acceptor; R: aromatic ring. ^bThe ranking is a measure of the rarity of pharmacophore models and how well they mapped to the training set compounds. ^cDirect hit and partial hit indicate whether a compound in a training set mapped to all the features of the hypothesis and just partially mapped to the pharmacophore model.

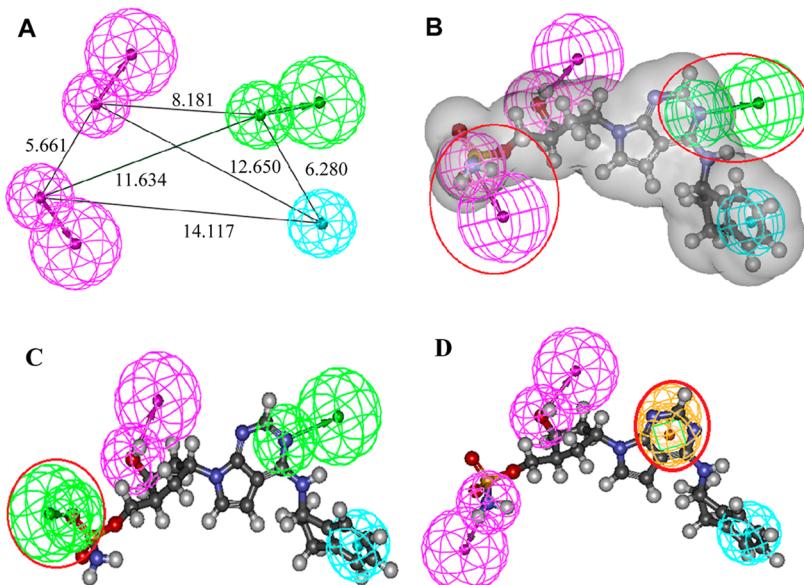


Figure 2. 3D arrangements of the best ligand-based pharmacophore model, Hypo1, and the MLN4924 mapping results of three representative hypotheses. The pharmacophore features are colored with green (hydrogen bond donor), orange (aromatic ring), and cyan (hydrophobic group). (A) The space organization of Hypo1. The distances between the feature centers are represented in solid black lines. (B) MLN4924 mapping result of Hypo1 with the shape constraint. (C) MLN4924 mapping result of Hypo3. (D) MLN4924 mapping result of Hypo4.

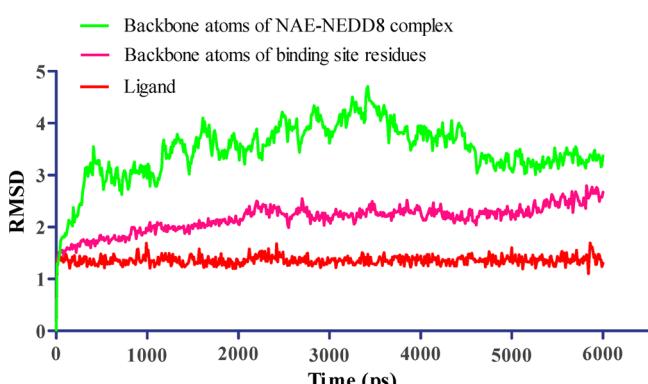


Figure 3. Plots reflecting the RMSD of the NAE-NEDD8 complex and ligand between the initial structure of the NAE complex and respective trajectory snapshots versus time.

conserved throughout the MD simulation. As shown in Figure 4A, after pharmacophore overlay several bunches of feature elements were found distributed around the binding site which contains some critical residues, such as Asp100, Ile148, Met101, Ile170, and Ala171 of NAE as well as Gly76 of NEED8. According to the residue each feature belongs to, total features were clustered to 8 sets which consist of 3 HBDs, 3 HBAs, and 2 hydrophobic features. In each set, the number of constituent elements which showed the frequency of a certain pharmacophore feature presented in different complexes was counted to determine the significance of every clustered feature. It is believed that features which have higher occurrence in complexes throughout the entire MD running are likely to be more important than those present only a few times. The statistical frequency of all clustered features and their corresponding residue partners were listed in Table 2.

One can see that seven of the eight pharmacophore features demonstrated a consistent appearance during the MD simulation with frequency more than 60%. The only exception was the HBA (A2) feature mapping to the backbone nitrogen

of Met101, which achieved a relatively lower frequency of 40%. Thus it was removed from our model to avoid redundancy. Among the remaining 7 features, four features exhibited the highest frequency of 90%. The A3 feature representing the HBA on the ligand side was mapped to the backbone nitrogen of Ile148. The feature of D2 stood for the HBD feature for putative ligands, corresponding to the carboxyl oxygen atom of Asp100 in the middle of the binding site. While the two hydrophobic subpockets surrounded by Met101 and Leu166 as well as Ile170, Ala171, and Trp174 were respectively expressed as the feature H1 and H2, which dictated the hydrophobic feature required in the cytosolic port. The D3 was the fifth remarkable feature which reflected the HBD whose partner was side chain oxygen atom of Gln149. The sixth frequently occurring feature was A1 which referred to the hydrogen bonding interaction between the HBA of ligand and Asp167. At last a consistent HBD element (D1) was found pointed to the reactive residue of Gly76 of NEED8, suggesting the close proximity between the sulfamoyl group of ligand and the terminal carbonyl group of NEED8, which facilitates the following reaction (Figure 4B).

For a complete pharmacophore construction based on protein structure, excluded volumes have to be included to reflect the steric restriction and inaccessible areas for potential ligands. In our study, 23 excluded volume features were added based on the positions of several highly conserved residues within a binding site, such as Ile170, Ala171, Ile148, Gln149, Leu166, Met101, Asp100, and Asp167 of NAE as well as Gly76 of NEED8. The complete pharmacophore model of Hypo8, including excluded volumes, was shown in Figure 4C.

The seven chemical features included in the dynamic SBP model would be too restrictive for small molecules and will cause the inefficiency of pharmacophore-based screening. Hence it is necessary to determine which features are indispensable for ligand binding and which features are optional. In order to make a judicious choice, the seven training set compounds used in the construction of LBP were

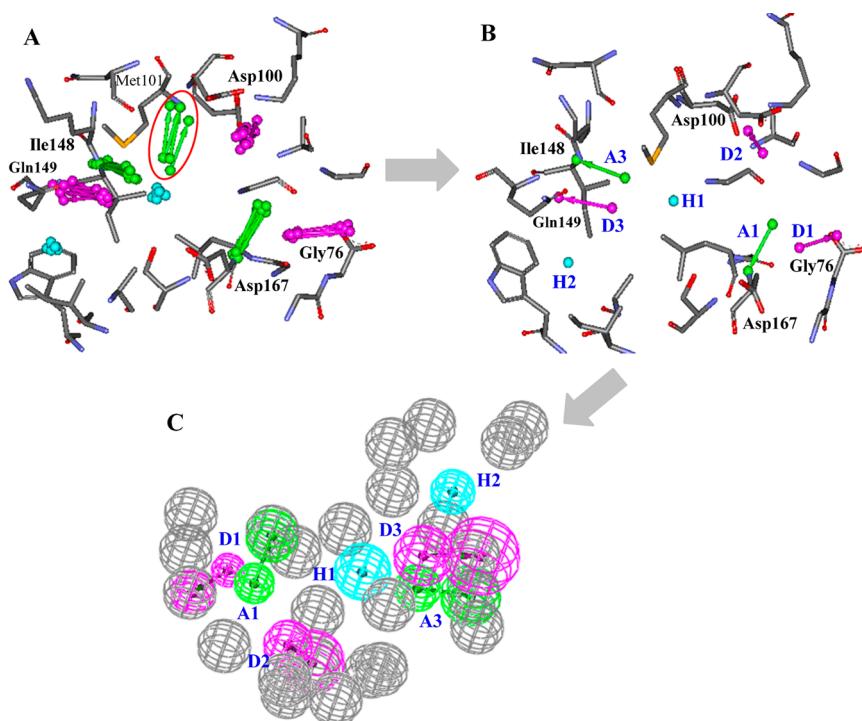


Figure 4. (A) Superimposition of all generated 10 pharmacophore models in the context of the NAE binding site. Purple represents the hydrogen bond donor; green is the hydrogen bond acceptor, and cyan stands for hydrophobic feature. (B) Average feature of each cluster emerging after feature clustering. (C) The final SBP model, HypoS, with excluded volumes.

Table 2. Statistical Frequency of Clustered Features and Their Acting Residues

no.	ID	feature type	frequency ^a	residue partner
1	D1	HBD	60%	Gly76 (NEDD8)
2	D2	HBD	90%	Asp100
3	D3	HBD	80%	Gln149
4	A1	HBA	70%	Asp167
5	A2	HBA	40%	Met101
6	A3	HBA	90%	Ile148
7	H1	hydrophobic	90%	Met101, Leu166
8	H2	hydrophobic	90%	Ile170, Ala171, Trp174

^aThe statistical frequency of a given feature is calculated as the ratio of its occurrence number in total 10 pharmacophore models.

docked to the model. The results showed that all seven compounds were successfully mapped to five of the seven features: D1, A1, D2, A2, and H2. However compound 2 did not match the feature of D3, and the H1 feature was also found missing in compounds 3–6 (Figure 5). Based on this analysis, the five features contained in all the training set ligands were set as a compulsory requirement in the SBP model, leaving the remaining two features, D3 and H1, as optional. The decision

to allow a partial match would be much more practical and preferable when a high constrained pharmacophore model was applied in a screening experiment.

3.3. Pharmacophore Model Validation. The two generated models were validated by screening an external decoy database containing 30 potent NAE inhibitors and 1000 drug-like noninhibitors with similar physical properties (Table S1). The enrichment factor and goodness of hit score were calculated for each pharmacophore model based on their performance in the library screening. As shown in Table 3, the

Table 3. Results of Pharmacophore Validation

parameter	Hypo1	Hypo1M	HypoS
total molecules in database (D)	1030	1030	1030
total number of actives in database (A)	30	30	30
total hits (Ht)	134	37	43
active hits (Ha)	28	27	28
% yield of actives $[(Ha/Ht) \times 100]$	20.90	72.97	65.12
% ratio of actives $[(Ha/A) \times 100]$	93.33	90.00	93.33
% false negatives rate $[(A-Ha)/A \times 100]$	6.67	10.00	6.67
% false positives rate $[(Ht-Ha)/(D-A) \times 100]$	10.60	1.00	1.50
enrichment factor (EF)	7.18	24.32	22.36
GH score	0.35	0.76	0.71

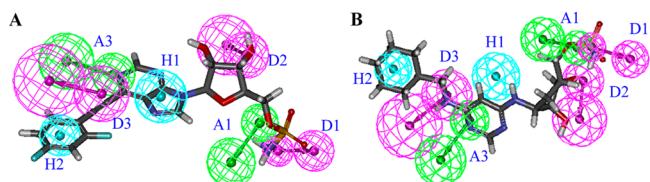


Figure 5. (A) HypoS mapped with compound 2. (B) HypoS superimposed with compound 4. Excluded volumes were not included in these figures for clarification purposes.

screening of HypoS identified 93.33% true positives and 1.5% decoys, and the overall EF and GH values of HypoS were 22.36 and 0.71, respectively, suggesting it was 22 times more probable to pick an active compound over an inactive one. On the other hand, Hypo1 also successfully discovered 28 actives in total 30 known NAE inhibitors but at the expense of higher number of decoys. The ratio of false positives for Hypo1 was 10.6%, almost ten times higher than that of HypoS, which made it less fitting in practical use. The reason for this poor performance

may lie in the limited feature constrains contained in the pharmacophore model, and there will be diverse sets of chemical entity large enough for the mapping of four features. Thus, with the aim to improve the selectivity of Hypo1, a shape constraint based on the crystallographycal structure of MLN4924 was added to the original model, which means a compound would be considered as a hit only if it can match both the pharmacophore model and shape constraint simultaneously (Figure 2B). As shown in Table 3, the modified model, Hypo1M, significantly reduced the ratio of false positives to 1.0% and exhibits an improved screening performance with the EF (24.32) and GH (0.76) values comparable to those of HypoS. Finally the validated HypoS and Hypo1 M were selected as queries for screening of large databases.

Interestingly, during the screening test, a recently reported noncovalent NAE inhibitor¹² in the decoy database was also identified by both of our pharmacophore models. It has scored a Fit value of 3.14 in the mapping with HypoS and was highly ranked by HypoS among the top 3% of screened database, and the interaction model proposed from the docking study¹² was also consistent with our pharmacophore mapping results (Figure S1). On the contrary, its mapping with Hypo1 M was relatively poor, resulting in a low Fit value of 0.24. The huge disparity between the two Fit values was quite predictable. Since Hypo1 M was built from the active conformations of covalent inhibitors, the space arrangement of feature elements would remarkably differentiate from that of noncovalent ligands which limited its scope of application for noncovalent inhibitors. On the other hand, HypoS was the result of pharmacophore overlay derived from a MD run. Before the MD simulation, the covalent linkage was broken, and pharmacophore features were extracted from a noncovalent interaction map complementing the hot spots of NAE active site. Thus, theoretically, HypoS has the capability to identify novel noncovalent ligands for NAE, and the identification and high rank of this noncovalent ligand by HypoS would just serve as strong evidence to this assumption.

3.4. Covalent Docking. Until now, the most potent class of NAE inhibitors is MLN4924 and its analogues. This series of ATP mimetic compounds exerts their inhibitory effect in a unique fashion which involves a covalent linkage between the sulfamoyl group and the Gly76 of NEDD8. A recent study has demonstrated the importance of this covalent bond by showing a significant activity loss when the interaction was blocked.¹⁸ Therefore, the sulfamoyl group was also retained in our VS campaign, and we attempted to search for a new chemical scaffold which could facilitate this reaction. In this case, the covalent docking module available in Gold 5.0 was adopted to carry out the docking-based screening. In comparison with other covalent docking programs, Gold applies a high efficient docking algorithm which allows a versatile and ligand-independent way to connect two reaction groups during the covalent docking process. This implemented algorithm could significantly save the computer cost for single molecule docking and render us a powerful tool for the VS of covalently linked ligands.⁴¹ To validate the reliability of the proposed docking application, its capabilities of pose prediction and actives discrimination were explored in both native docking and decoy database screening scenarios using the covalently bonded NAE structure (3GZN).

In the native docking experiment, four types of docking score, Goldscore, Chemscore, ASP and CHEMPLP, available in

Gold, were independently investigated for their abilities to identify correct binding poses. The docking accuracy was determined by calculating the RMSD of heavy atoms between the top-ranked poses and the crystallographic counterpart. According to the results shown in Table 4, among the top-

Table 4. Native Docking Results Using Four Types of Docking Score

scoring function	RMSD (Å)
Goldscore	0.78
Chemscore	9.05
ASP	0.57
CHEMPLP	0.79

ranked poses, ASP, Goldscore, and CHEMPLP were able to reproduce the crystallographic pose of MLN4924 within an acceptable tolerance (RMSD < 2.0 Å), while docking with Chemscore exhibited a poor performance in pose prediction by showing high a RMSD value of 9.05. For this reason, Goldscore, ASP, and CHEMPLP were selected for further evaluation.

After that, a decoy data set comprising 30 covalent NAE inhibitors and 1000 sulfamoyl-contained drug-like decoys was used to assess the ability of the three remaining docking scores to discriminate actives from inactives. The overall screening performance of the three scores was revealed in the receiver operating characteristic (ROC) curve. In one extreme situation, an ideal scoring function which ranks all the actives over any decoys would have the largest area under the curve (AUC), while in the other extreme, the ROC curve of a randomly ranking scoring function would be a diagonal in which the AUC value is 0.5. Referring to Figure 6, apparently, the Goldscore

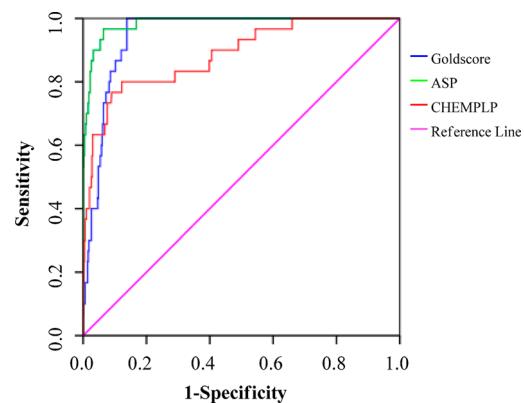
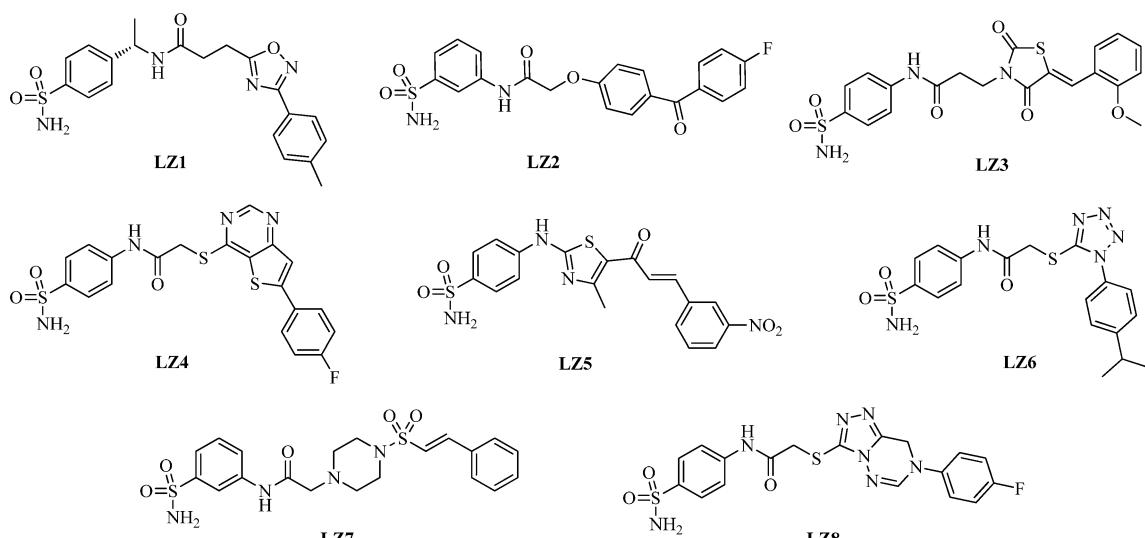


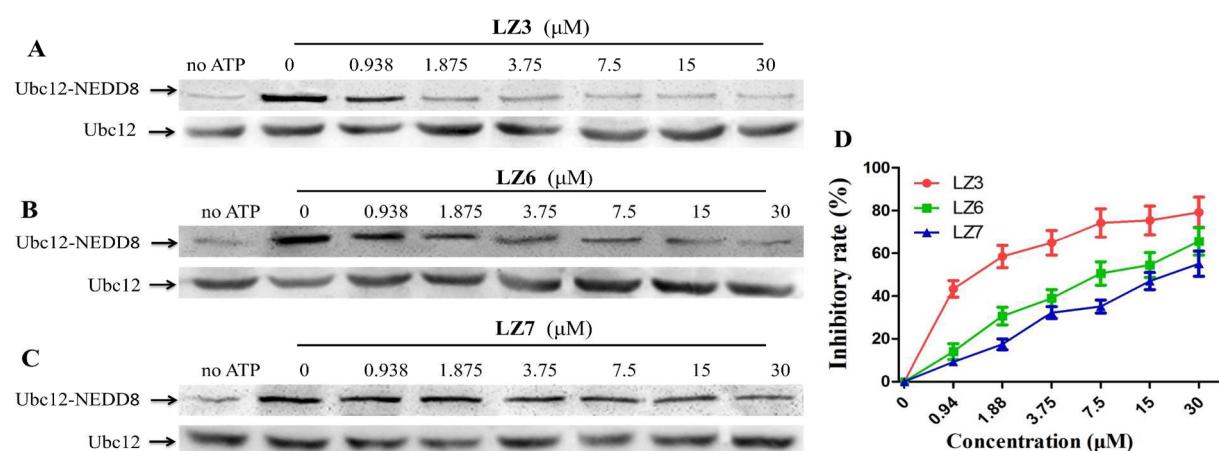
Figure 6. ROC curve for the three scoring functions. Shown in blue is the screening performance of Goldscore; shown in green is the screening performance of ASP, and the red line represents the performance of CHEMPLP.

and ASP significantly outperformed CHEMPLP in the ranking capability by showing a higher AUC value. Thus Goldscore and ASP were chosen for the consensus ranking of screening hits. Besides, another essential function of ROC analysis is to decide the optimal threshold which strikes efficient balance between selectivity and specificity. In this term, the values of 20.58 in ASP score and 25.93 in Goldscore were settled because they would truncate the overall false positive ratio down to 15% and meanwhile only sacrificing a small fraction of true positives.

**Figure 7.** Chemical structure of tested compounds**Table 5. General Results of Library Screening and Biological Tests**

name	ZINC ID	HypoS	Hypo1M	Goldscore	ASP	inhibitory rate ^a (%)	$\text{IC}_{50}^b\ (\mu\text{M})$
LZ3	ZINC13573581	3.30	2.57	26.70	24.17	79.23 ± 4.76	1.06 ± 0.18
LZ6	ZINC10382178	1.99	3.42	32.37	26.25	65.59 ± 3.77	9.18 ± 1.25
LZ7	ZINC12550064	4.01	1.89	34.35	20.30	55.16 ± 3.93	19.71 ± 2.43
MLN4924		5.31	4.00	45.82	39.06	97.37 ± 2.85	11.87 ± 2.06^c

^aInhibitory activity of compounds against NAE *in vitro* was determined with a concentration of 30 μM by Western blot analysis. ^b IC_{50} , compound concentration required to inhibit the expression of Ubc12-NEDD8 and Ubc12 by 50%. ^cThe IC_{50} value is expressed in nM. Data are expressed as the mean \pm SEM from at least three independent experiments performed in triplicate.

**Figure 8.** LZ3, LZ6, and LZ7 inhibit Ubc12-NEDD8 formation in a dose-dependent manner. (A), (B), and (C) show that NAE charged with Ubc12 and NEDD8 was incubated for 2 h with indicated concentrations of tested compounds. The reaction mixture was fractionated by nonreducing SDS-PAGE and immunoblotted using Ubc12 antibody. (D) Densitometry analysis of the Western blot showing compounds-mediated reduction of Ubc12-NEDD8.

3.5. Library Screening and Biological Test Results. The screening of a commercial database has provided a fruitful resource for lead identification. In this research, with the aim of identification of novel covalent inhibitors of NAE, a focused library derived from the ZINC database was constructed in which all compounds contained free sulfamoyl groups. Then database screening against this focused library was carried out with the two verified pharmacophore models, which resulted in 241 and 48 hit compounds from Hypo1 M and HypoS, respectively. After removal of duplicates, a total of 256 hit compounds were covalently docked to the active site of NAE,

and 64 hits with the ASP score above 20.58 or Goldscore larger than 25.93 were extracted for visual inspection. Molecules which can form a favorable hydrogen-bonding interaction with key residues like Asp100, Ile148, and Asp167 were retained, while those with high strained binding conformations were discarded. Ultimately, eight novel compounds with various scaffolds were selected and purchased from commercial suppliers for biological tests (Figure 7).

All selected compounds were first evaluated for their inhibitory activity against NAE by a cell-free assay, and MLN4924 was used as a positive control. NAE was incubated

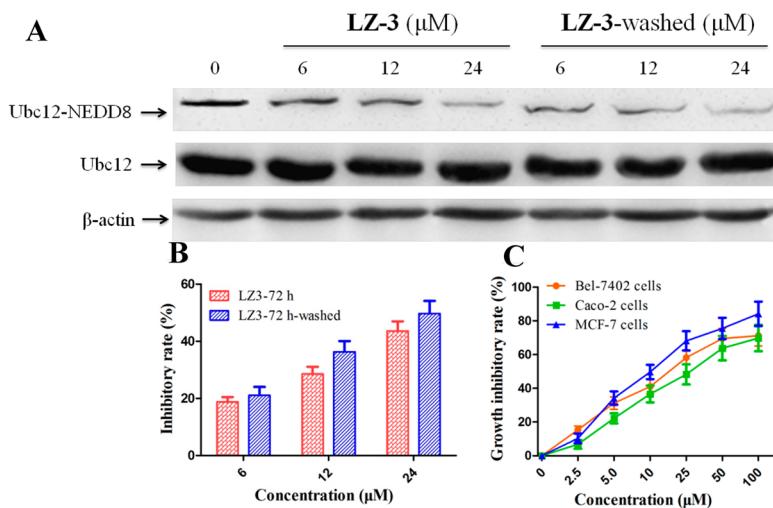


Figure 9. (A) The inhibitory activity of LZ3 remained after the washout experiment. MCF-7 cells were treated with various concentrations of LZ-3 for 72 h to induce senescence. One part of the cells was lysed, and the other part was washed and cultured in drug-free medium for an additional 24 h (LZ-3-washed); the expression of Ubc12-NEDD8 and Ubc12 was determined by Western blot analysis. Expression of β -actin was used as an internal control. (B) The inhibitory rate of NAE in MCF-7 cells. Data are expressed as the mean \pm SEM from at least three independent experiments performed in triplicate. (C) LZ3 concentration-dependent inhibited the proliferation of Bel-7402, MCF-7, and Caco-2 cells *in vitro*. Cells were treated with LZ3 (2.5–100 μM) for 7 days, respectively. The cytotoxicity of LZ3 was analyzed by MTT assay. Data are expressed as the mean \pm SEM from at least three independent experiments performed in triplicate.

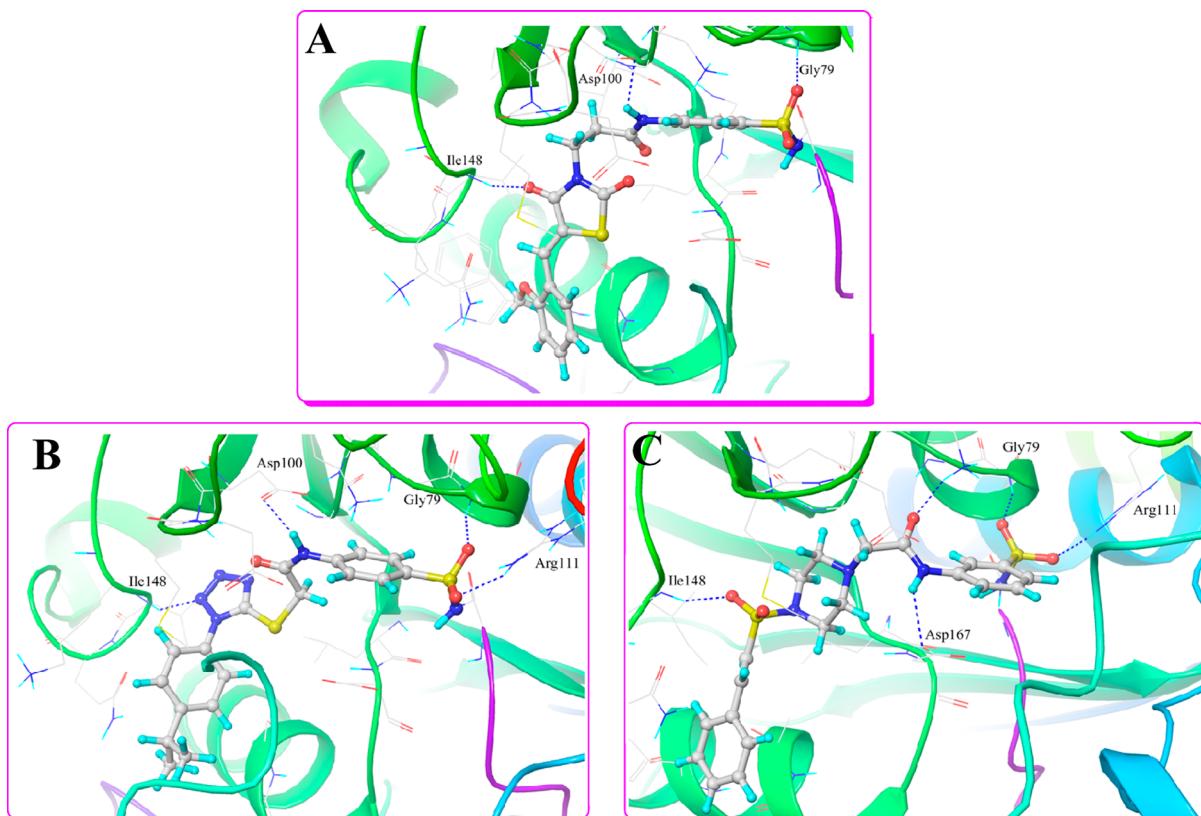


Figure 10. Predicted binding modes of (A) LZ3, (B) LZ6, and (C) LZ7 to NAE. The interaction modes are viewed in PyMol. Proteins are displayed in cartoon with binding residues shown in lines. Ligands are presented in sticks.

with NEDD8 and Ubc12 in the presence of 30 μM compounds or vehicle for 2 h. NAE promotes NEDDylation of Ubc12, leading to the formation of the Ubc12-NEDD8 thioester product.⁷ Table 5 summarized the results of overall VS experiment and the corresponding biological tests. Among the eight tested hits, three compounds, LZ3, LZ6, and LZ7,

displayed remarkable inhibition against this process with the inhibitory rate above 55%. Then their IC₅₀ values were further determined by Western blot analysis (Table 5, Figure 8). Notably, LZ3 demonstrated the highest inhibitory activity with an IC₅₀ value of $1.06 \pm 0.18 \mu\text{M}$ which is almost 10–20-fold more potent than LZ6 and LZ7 (IC₅₀ of $9.18 \pm 1.25 \mu\text{M}$ and

19.71 ± 2.43 μM, respectively). It was observed that the two compounds **LZ3** and **LZ6** with the highest inhibitory activity both possess a common para-sulfamoyl-substituted phenylamino group in their structures. This particular moiety not only contained the required sulfamoyl group ready for chemical linkage but also enabled a perfect space arrangement which allows the formation of an important hydrogen bond between the phenylamino group and Asp 100.

Since NAE inhibitors are reported as effective suppressors of tumor cell growth by causing cell cycle arrest and apoptosis,^{42,43} the most potent hit, **LZ3**, was then further assessed for its efficacy to inhibit the proliferation of several cancer cell lines, including Caco-2, MCF-7 and Bcl-7402 cells. The data obtained from MTT assay showed that **LZ3** displayed potent growth inhibition in all tested cell lines with IC₅₀ values ranging from 12.3 to 29.5 μM (Figure 9C). Subsequent Western blot analysis of the MCF-7 cell lysis revealed a concentration-dependent reduction of Ubc12-NEDD8 formation by **LZ3**, which served as a solid proof for its inhibitory efficacy in cells (Figure 9A and 9B).

To address the question that whether the covalent binding of **LZ3** was formed, a cell-based washout experiment^{44–46} was performed. Briefly, after treatment with **LZ3** for three days, MCF-7 cells were washed by PBS to remove the unbounded compound followed by incubation in fresh medium for another 24 h. Normally, under this condition, a noncovalent ligand would dissociate quickly from protein, leading to the recovery of enzyme activity, whereas a covalent binder would remain its inhibitory effect, keeping the recovery of enzyme activity in check. In our study, as expected, no recovery of NAE activity was observed after washout of the inactivated enzyme by **LZ3**, thus suggesting an occurring covalent binding event (Figure 9A and 9B).

In order to gain further detailed insight into the interaction mechanism of the three hit compounds, their binding modes with NAE were predicted by covalent docking. As shown in Figure 10, all three compounds adopted a similar binding pattern as MLN4924. In addition to the covalent binding, **LZ3** and **LZ6** interact with Asp100 through hydrogen bond interaction between their phenylamino group and the residue side chain. This specific contact in NAE binding is extremely crucial, because the absence of this interaction feature would give rise to a dramatic decrease in the binding affinity which was directly demonstrated by the case of **LZ7**. In the structure of **LZ7**, the sulfamoyl group is at the meta-position of phenylamino group which shortens the distance between the sulfamoyl and amino nitrogen when compared to its para-substituted counterpart in **LZ3** and **LZ6**. It is this curtailed distance which prevents molecule from taking an appropriate conformation to formulate the proposed H-bond interaction with Asp100. Thus the binding affinity of **LZ7** has been shown inferior to the other two hits. In addition, all three hits form a hydrogen bond with Ile148 and make a favorable interaction with the hydrophobic pocket in the cytosolic region using diverse scaffolds: **LZ3** employs a thiazolidine-2,4-dione ring which contacts Ile149 with its carbonyl and the 2-methoxybenzene ring extended into the hydrophobic region to develop a desirable hydrophobic interaction; **LZ6** makes the hydrogen bonding with Ile148 through the nitrogen atom in its tetrazole ring, and similarly a 4-isopropylbenzene moiety was stretched into the cytosolic pocket; while in the case of **LZ7**, the hydrogen bonding interaction is achieved by another sulfonyl group in the structure, and the terminal alkenyl benzene ring

contributes to the hydrophobic interaction between NAE and the ligand. However, despite these common interaction features they shared, the three compounds also present several characteristic interaction modes. For instance, compound **LZ7**, although missing the hydrogen bond with Asp100, engages in a hydrogen bonding network involving interaction with Asp167 and Lys124. This type of interaction might compensate for the affinity loss due to the absence of Asp 100 interaction.

These three identified hits presented distinct structures for the known NAE inhibitors ensemble. The novelty of these three scaffolds was evaluated by pairwisely calculating the Tanimoto similarity indices between hit compounds and known NAE inhibitors in the ChEMBL database. The Find Similar Molecules by Fingerprints protocol combined with FCFP_6 fingerprints in DS was used to carry out the calculation. The results showed that all three compounds exhibited low Tanimoto similarity indices with respect to known NAE inhibitors (0.10–0.15 for **LZ3**, 0.09–0.10 for **LZ6**, and 0.07–0.10 for **LZ7**), thus they can be considered as novel classes of NAE inhibitors. It is noted that these compounds are still at the first stage of drug development, and they have not been subjected to structure modification and optimization. There is a possibility of improved bioactivity when careful optimization steps are taken. Therefore, all these above have proved that our combined VS strategy is proper for the identification of novel covalent NAE inhibitors.

4. CONCLUSIONS

Recently, fueled by the great success of covalent kinase inhibitors,^{47,48} there is a growing enthusiasm to develop other targeted covalent drugs right from the start.^{49,50} However, the discovery of covalent ligands by VS has been traditionally deemed as a restricted area because most of the currently available VS technologies are designed to account for the noncovalent interaction between ligands and proteins. In recent research, Schröder and his team have demonstrated the first successful VS application of covalent docking in identifying potential irreversible human cathepsin K inhibitors.⁵¹ Different from that VS strategy where only covalent docking was involved, in the present study, we described a new combined VS strategy which incorporated both ligand- and dynamic structure-based pharmacophore modeling as well as covalent docking to discover novel covalent NAE inhibitors. Finally, through this hierarchical screening procedure, three compounds were discovered as new classes of covalent NAE inhibitors. Among them, **LZ3** has been shown to possess the most preeminent inhibitory potency against NAE in a cell-free system with an IC₅₀ value of 1.06 ± 0.18 μM. Further experiment of **LZ3** also confirmed its efficacy in a cellular system, and its covalent binding mechanism was proved by a cell-based washout experiment which showed no recovery of enzyme activity after PBS dilution. Although the bioactivity of these identified hit compounds is in the moderate micromolar range, the structural novelty and diversity of these hits would render us a wide range of optimization space which might serve as a starting point for the development of new types of potent NAE inhibitors.

■ ASSOCIATED CONTENT

● Supporting Information

Chemical structure of reported noncovalent NAE inhibitor and its alignment with HypoS, NAE inhibitory activity of all eight

hits in a cell-free system, physical properties of decoy molecules. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Authors

*Phone/Fax: +86-25-86205849. E-mail: huijicpu@163.com (H.J.).

*Phone/Fax: +86-25-83271015. E-mail: yslai@cpu.edu.cn (Y.L.).

Author Contributions

[†]These two authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We sincerely thank Dr. Yadong Chen for his professional advices in drafting the manuscript. This work was supported by the Project Program of State Key Laboratory of Natural Medicines, China Pharmaceutical University (No. ZJ12020), the Natural Science Foundation of Jiangsu Province (No. BK2011626), the College Students Innovation Project for the R&D of Novel Drugs (No. J1030830), and the National College Students Innovation Experiment Program (No. G13056).

REFERENCES

- (1) Cohen, F. E.; Kelly, J. W. Therapeutic approaches to protein-misfolding diseases. *Nature* **2003**, *426*, 905–909.
- (2) Balch, W. E.; Morimoto, R. I.; Dillin, A.; Kelly, J. W. Adapting proteostasis for disease intervention. *Science* **2008**, *319*, 916–919.
- (3) Mu, T. W.; Ong, D. S. T.; Wang, Y. J.; Balch, W. E.; Yates, J. R., III; Segatori, L.; Kelly, J. W. Chemical and biological approaches synergize to ameliorate protein-folding diseases. *Cell* **2008**, *134*, 769–781.
- (4) Hershko, A. The ubiquitin system for protein degradation and some of its roles in the control of the cell division cycle. *Cell Death Differ.* **2005**, *12*, 1191–1197.
- (5) Kane, R. C.; Bross, P. F.; Farrell, A. T.; Pazdur, R. Velcade®: U.S. FDA approval for the treatment of multiple myeloma progressing on prior therapy. *Oncologist* **2003**, *8*, 508–513.
- (6) Kane, R. C.; Dagher, R.; Farrell, A.; Ko, C. W.; Sridhara, R.; Justice, R.; Pazdur, R. Bortezomib for the treatment of mantle cell lymphoma. *Clin. Cancer Res.* **2007**, *13*, 5291–5294.
- (7) Soucy, T. A.; Smith, P. G.; Rolfe, M. Targeting NEDD8-activated cullin-RING ligases for the treatment of cancer. *Clin. Cancer Res.* **2009**, *15*, 3912–3916.
- (8) Milhollen, M. A.; Traore, T.; Adams-Duffy, J.; Thomas, M. P.; Berger, A. J.; Dang, L.; Dick, L. R.; Garnsey, J. J.; Koenig, E.; Langston, S. P.; Manfredi, M.; Narayanan, U.; Rolfe, M.; Staudt, L. M.; Soucy, T. A.; Yu, J.; Zhang, J.; Bolen, J. B.; Smith, P. G. MLN4924, a NEDD8-activating enzyme inhibitor, is active in diffuse large B-cell lymphoma models: rationale for treatment of NF-κB-dependent lymphoma. *Blood* **2010**, *116*, 1515–1523.
- (9) Chairatvit, K.; Ngamkitidechakul, C. Control of cell proliferation via elevated NEDD8 conjugation in oral squamous cell carcinoma. *Mol. Cell. Biochem.* **2007**, *306*, 163–169.
- (10) Lukkarila, J. L.; da Silva, S. R.; Ali, M.; Shahani, V. M.; Xu, G. W.; Berman, J.; Roughton, A.; Dhe-Paganon, S.; Schimmer, A. D.; Gunning, P. T. Identification of NAE inhibitors exhibiting potent activity in leukemia cells: exploring the structural determinants of NAE specificity. *ACS Med. Chem. Lett.* **2011**, *2*, 577–582.
- (11) Leung, C. H.; Chan, D. S. H.; Yang, H.; Abagyan, R.; Lee, S. M.; Zhu, G. Y.; Fong, W. F.; Ma, D. L. A natural product-like inhibitor of NEDD8-activating enzyme. *Chem. Commun.* **2011**, *47*, 2511–2513.
- (12) Zhong, H. J.; Ma, V. P. Y.; Cheng, Z.; Chan, D. S. H.; He, H. Z.; Leung, K. H.; Ma, D. L.; Leung, C. H. Discovery of a natural product inhibitor targeting protein NEDDylation by structure-based virtual screening. *Biochimie* **2012**, *94*, 2457–2460.
- (13) Nawrocki, S. T.; Griffin, P.; Kelly, K. R.; Carew, J. S. MLN4924: a novel first-in-class inhibitor of NEDD8-activating enzyme for cancer therapy. *Expert Opin. Invest. Drugs* **2012**, *21*, 1563–1573.
- (14) Soucy, T. A.; Smith, P. G.; Milhollen, M. A.; Berger, A. J.; Gavin, J. M.; Adhikari, S.; Brownell, J. E.; Burke, K. E.; Cardin, D. P.; Critchley, S. An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **2009**, *458*, 732–736.
- (15) Brownell, J. E.; Sintchak, M. D.; Gavin, J. M.; Liao, H.; Buzzese, F. J.; Bump, N. J.; Soucy, T. A.; Milhollen, M. A.; Yang, X.; Burkhardt, A. L. Substrate-assisted inhibition of ubiquitin-like protein-activating enzymes: the NEDD8 E1 inhibitor MLN4924 forms a NEDD8-AMP mimetic in situ. *Mol. Cell* **2010**, *37*, 102–111.
- (16) Milhollen, M. A.; Thomas, M. P.; Narayanan, U.; Traore, T.; Riceberg, J.; Amidon, B. S.; Bence, N. F.; Bolen, J. B.; Brownell, J.; Dick, L. R.; Loke, H. K.; McDonald, A. A.; Ma, J. Y.; Manfredi, M. G.; Sells, T. B.; Sintchak, M. D.; Yang, X. F.; Xu, Q.; Koenig, E. M.; Gavin, J. M.; Smith, P. G. Treatment-emergent mutations in NAE beta confer resistance to the NEDD8-activating enzyme inhibitor MLN4924. *Cancer Cell* **2012**, *21*, 388–401.
- (17) Maluf, F. V.; Andricopulo, A. D.; Oliva, G.; Guido, R. V. A pharmacophore-based virtual screening approach for the discovery of Trypanosoma cruzi GAPDH inhibitors. *Future Med. Chem.* **2013**, *5*, 2019–2035.
- (18) Sakkiah, S.; Thangapandian, S.; John, S.; Lee, K. W. Pharmacophore based virtual screening, molecular docking studies to design potent heat shock protein 90 inhibitors. *Eur. J. Med. Chem.* **2011**, *46*, 2937–2947.
- (19) Distinto, S.; Esposito, F.; Kirchmair, J.; Cardia, M. C.; Gaspari, M.; Maccioni, E.; Alcaro, S.; Markt, P.; Wolber, G.; Zinzula, L.; Tramontano, E. Identification of HIV-1 reverse transcriptase dual inhibitors by a combined shape-, 2D-fingerprint- and pharmacophore-based virtual screening approach. *Eur. J. Med. Chem.* **2012**, *50*, 216–229.
- (20) Waltenberger, B.; Wiechmann, K.; Bauer, J.; Markt, P.; Noha, S. M.; Wolber, G.; Rollinger, J. M.; Werz, O.; Schuster, D.; Stuppner, H. Pharmacophore modeling and virtual screening for novel acidic inhibitors of microsomal prostaglandin E2 synthase-1 (mPGES-1). *J. Med. Chem.* **2011**, *54*, 3163–3174.
- (21) Critchley, S.; Gant, T. G.; Langston, S. P.; Olhava, E. J.; Peluso, S. Inhibitors of E1 activating enzymes. U.S. Patent US 20060189636, Aug. 24, 2006; *Chem. Abstr.* **2006**, *145*, 230849.
- (22) Claiborne, C. F.; Critchley, S.; Cullis, C.; Langston, S. P.; Mizutani, H.; Olhava, E. J.; Peluso, S.; Visiers, I.; Vyskocil, S.; Weatherhead, G. S. Heteroaryl compounds useful as inhibitors of E1 activating enzymes. U.S. Patent US 20120071482, Feb. 14, 2008; *Chem. Abstr.* **2008**, *148*, 262855.
- (23) Wang, Z.; Smith, E. D.; Veal, J.; Huang, K. H.; Atkinson, R. N.; Jiang, R. Sulfonamide, sulfamate, and sulfamothioate derivatives. U.S. Patent US 20120077814, Mar. 29, 2012; *Chem. Abstr.* **2012**, *156*, 477751.
- (24) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (25) Clement, O. O.; Mehl, A. T. HipHop: pharmacophore based on multiple common-feature alignments. In *Pharmacophore perception, development, and use in drug design*; Güner, O. F., Ed.; International University Line: La Jolla, CA, 2000; pp 69–84.
- (26) Ryckaert, J. P.; Ciccotti, 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.
- (27) York, D. M.; Darden, T. A.; Pedersen, L. G. The effect of long-range electrostatic interactions in simulations of macromolecular

- crystals: A comparison of the Ewald and truncated list methods. *J. Chem. Phys.* **1993**, *99*, 8345–8348.
- (28) Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- (29) Böhm, H. J. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 61–78.
- (30) Huang, D.; Gu, Q.; Ge, H.; Ye, J.; Salam, N. K.; Hagler, A.; Chen, H. Z.; Xu, J. On the value of homology models for virtual screening: discovering hCXCR3 antagonists by pharmacophore-based and structure-based approaches. *J. Chem. Inf. Model.* **2012**, *52*, 1356–1366.
- (31) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49*, 6789–6801.
- (32) 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.
- (33) Martin, E. J.; Blaney, J. M.; Siani, M. A.; Spellmeyer, D. C.; Wong, A. K.; Moos, W. H. Measuring diversity: experimental design of combinatorial libraries for drug discovery. *J. Med. Chem.* **1995**, *38*, 1431–1436.
- (34) Rogers, D.; Brown, R. D.; Hahn, M. Using extended-connectivity fingerprints with Laplacian-modified Bayesian analysis in high-throughput screening follow-up. *J. Biomol. Screening* **2005**, *10*, 682–686.
- (35) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins* **2003**, *52*, 609–623.
- (36) Schwartz, P. A.; Kuzmic, P.; Solowiej, J.; Bergqvist, S.; Bolanos, B.; Almaden, C.; Nagata, A.; Ryan, K.; Feng, J.; Dalvie, D.; Kath, J. C.; Xu, M.; Wani, R.; Murray, B. W. Covalent EGFR inhibitor analysis reveals importance of reversible interactions to potency and mechanisms of drug resistance. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 173–178.
- (37) Meagher, K. L.; Carlson, H. A. Incorporating protein flexibility in structure-based drug discovery: using HIV-1 protease as a test case. *J. Am. Chem. Soc.* **2004**, *126*, 13276–13281.
- (38) Carlson, H. A.; Masukawa, K. M.; Rubins, K.; Bushman, F. D.; Jorgensen, W. L.; Lins, R. D.; Briggs, J. M.; McCammon, J. A. Developing a dynamic pharmacophore model for HIV-1 integrase. *J. Med. Chem.* **2000**, *43*, 2100–2014.
- (39) Bowman, A. L.; Makriyannis, A. Approximating protein flexibility through dynamic pharmacophore models: application to fatty acid amide hydrolase (FAAH). *J. Chem. Inf. Model.* **2011**, *51*, 3247–3253.
- (40) Bowman, A. L.; Lerner, M. G.; Carlson, H. A. Protein flexibility and species specificity in structure-based drug discovery: dihydrofolate reductase as a test system. *J. Am. Chem. Soc.* **2007**, *129*, 3634–3640.
- (41) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (42) Swords, R. T.; Kelly, K. R.; Smith, P. G.; Garnsey, J. J.; Mahalingam, D.; Medina, E.; Oberheu, K.; Padmanabhan, S.; O'Dwyer, M.; Nawrocki, S. T.; Giles, F. J.; Carew, J. S. Inhibition of NEDD8-activating enzyme: a novel approach for the treatment of acute myeloid leukemia. *Blood* **2010**, *115*, 3796–3800.
- (43) Jia, L.; Li, H.; Sun, Y. Induction of p21-dependent senescence by an NAE inhibitor, MLN4924, as a mechanism of growth suppression. *Neoplasia* **2011**, *13*, 561–569.
- (44) Chatterjee, S.; Ator, M. A.; Bozyczko-Coyne, D.; Josef, K.; Wells, G.; Tripathy, R.; Iqbal, M.; Bihovsky, R.; Senadhi, S. E.; Mallya, S. Synthesis and biological activity of a series of potent fluoromethyl ketone inhibitors of recombinant human calpain I. *J. Med. Chem.* **1997**, *40*, 3820–3828.
- (45) Li, Y. B.; Wang, Z. Q.; Yan, X.; Chen, M. W.; Bao, J. L.; Wu, G. S.; Ge, Z. M.; Zhou, D. M.; Wang, Y. T.; Li, R. T. IC-4, a new irreversible EGFR inhibitor, exhibits prominent anti-tumor and anti-angiogenesis activities. *Cancer Lett.* **2013**, *340*, 88–96.
- (46) Xu, S.; Xu, T.; Zhang, L.; Zhang, Z.; Luo, J.; Liu, Y.; Lu, X.; Tu, Z.; Ren, X.; Ding, K. Design, synthesis, and biological evaluation of 2-oxo-3,4-dihydropyrimido[4,5-d]pyrimidinyl derivatives as new irreversible epidermal growth factor receptor inhibitors with improved pharmacokinetic properties. *J. Med. Chem.* **2013**, *56*, 8803–8813.
- (47) Sanderson, K. Irreversible kinase inhibitors gain traction. *Nat. Rev. Drug. Discovery* **2013**, *12*, 649–651.
- (48) Singh, J.; Petter, R. C.; Kluge, A. F. Targeted covalent drugs of the kinase family. *Curr. Opin. Chem. Biol.* **2010**, *14*, 475–480.
- (49) Hagel, M.; Niu, D.; St Martin, T.; Sheets, M. P.; Qiao, L.; Bernard, H.; Karp, R. M.; Zhu, Z.; Labenski, M. T.; Chaturvedi, P.; Nacht, M.; Westlin, W. F.; Petter, R. C.; Singh, J. Selective irreversible inhibition of a protease by targeting a noncatalytic cysteine. *Nat. Chem. Biol.* **2011**, *7*, 22–24.
- (50) Johnson, D. S.; Weerapana, E.; Cravatt, B. F. Strategies for discovering and derisking covalent, irreversible enzyme inhibitors. *Future Med. Chem.* **2010**, *2*, 949–964.
- (51) Schröder, J.; Klinger, A.; Oellien, F.; Marhöfer, R. J.; Duszenko, M.; Selzer, P. M. Docking-based virtual screening of covalently binding ligands: an orthogonal lead discovery approach. *J. Med. Chem.* **2013**, *56*, 1478–1490.