Supporting Information:

**Development and Validation of An Automated DNA-encoded Library Screening Data Analysis Platform: PB-DEL Auto Screening Analysis (PB-DELASA)**

*Keke Dong* †, *Xiangfei Meng* †, *Hongyi Diao* ‡, *Bing Qi* †, *Zhuangzhi Chen* †,*Wei Ma* †, *Yihang Zhang*†, *Minmin Yang* †, *Jing Zhao* \*,‡ *and Liu Liu* \*,†

**­**

†PharmaBlock Sciences (Nanjing), Inc., 81 Huasheng Road, Jiangbei New Area, Nanjing, Jiangsu, 210032, China.

‡State Key Laboratory of Coordination Chemistry, Chemistry and Biomedicine Innovation Center (ChemBIC), ChemBioMed Interdisciplinary Research Center at Nanjing University, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, 210093, PR China.

\* Correspondence to: lljxh99@gmail.com

The Supporting Materials file includes:

**Figure S1.** Comprehensive evaluation of the machine learning model performance. (a) Feature importance ranking showing Line\_SumCount as the most predictive feature. (b) Ten-fold cross-validation results illustrating consistent performance metrics across different data partitions. (c) Precision-Recall curve demonstrating exceptional classification performance (PR-AUC = 0.988) significantly above the random baseline

**Figure S2:** (a) Structural alignment diagram comparing the molecular docking conformation of AZD4573 (carbon atoms of the ligand colored by green) complexed with CDK9 with the crystal structure conformation (PDB: 6Z45; carbon atoms of the ligand colored by yellow). (b) 2D representation of hydrogen bonds and hydrophobic interactions within the docking conformation of CDK9-AZD4573 complex (figures were prepared using LIGPLOT). Dashed lines represent hydrogen bonds, while spiked residues forming hydrophobic interactions. (c) Overlay of docking conformation for AZD4573 and Compounds 1, 2, and 5. All four compounds bind to the same pocket of CDK9. Carbon atoms of AZD4573 are shown in green, compound 1 in white, compound 2 in blue, and compound 5 in magenta.

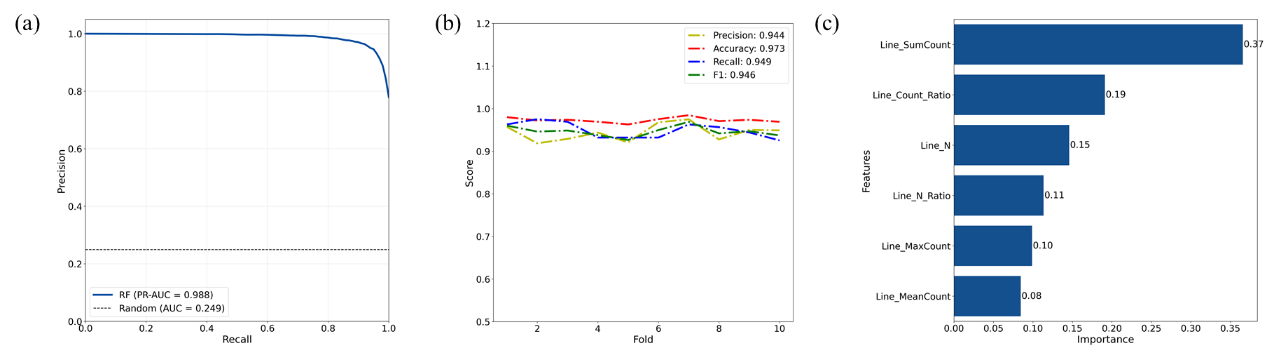
**Figure S3:** 2D representation of hydrogen bonds and hydrophobic interactions. Dashed lines indicate hydrogen bonds, while spiked residues form hydrophobic interactions with CDK9 complexes: a, 1; b: 2; c: 5 (figures were prepared using LIGPLOT).

**Figure S4:** Time-dependent RMSD of CDK9/Inhibitor complexes during MD simulation. The plot shows the RMSDs of the three inhibitors relative to their initial minimized structures over the course of MD simulations for the three CDK9/inhibitor complexes.

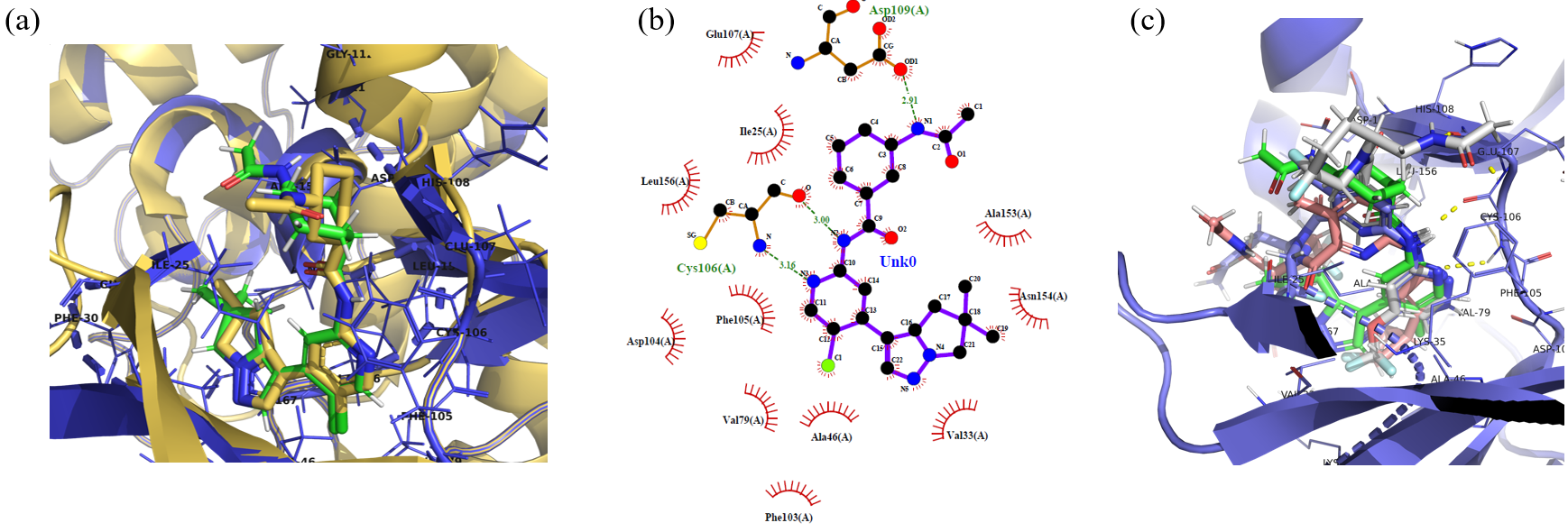
**Figure S5:** Per-Residue decomposition of ∆G for protein−inhibitor complexes. The graphs illustrate the decomposition of binding free energy (∆G) on a per-residue basis for complexes between CDK9 and (a), CDK9/1; (b), CDK9/2; (c), CDK9/5.

**Table S3:** Binding free energy components for the protein–inhibitor complexes calculated using MM-PBSA.

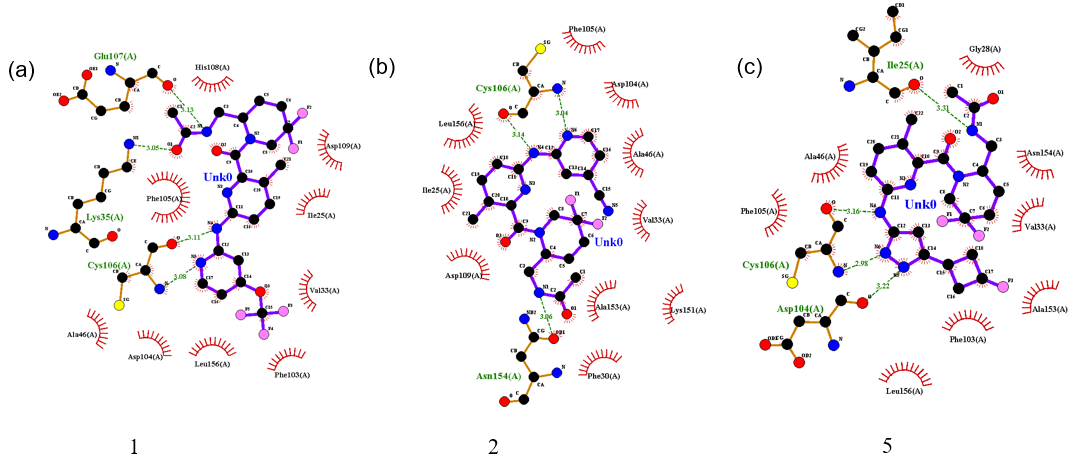
**Table S4:** Hydrogen bonds formed between inhibitors and CDK9 during MD simulations (The latter 20 ns) (Occupancy ≥ 50%)



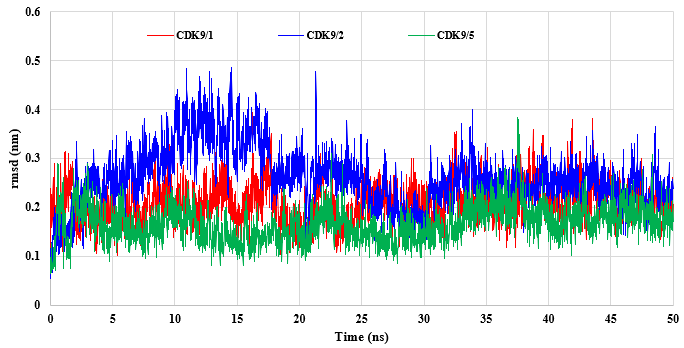
**Figure S1.** Comprehensive evaluation of the machine learning model performance. (a) Feature importance ranking showing Line\_SumCount as the most predictive feature. (b) Ten-fold cross-validation results illustrating consistent performance metrics across different data partitions. (c) Precision-Recall curve demonstrating exceptional classification performance (PR-AUC = 0.988) significantly above the random baseline.



**Figure S2.** (a) Structural alignment diagram comparing the molecular docking conformation of AZD4573 (carbon atoms of the ligand colored by green) complexed with CDK9 with the crystal structure conformation (PDB: 6Z45; carbon atoms of the ligand colored by yellow). (b) 2D representation of hydrogen bonds and hydrophobic interactions within the docking conformation of CDK9-AZD4573 complex (figures were prepared using LIGPLOT). Dashed lines represent hydrogen bonds, while spiked residues forming hydrophobic interactions. (c) Overlay of docking conformation for AZD4573 and Compounds 1, 2, and 5. All four compounds bind to the same pocket of CDK9. Carbon atoms of AZD4573 are shown in green, compound 1 in white, compound 2 in blue, and compound 5 in magenta.

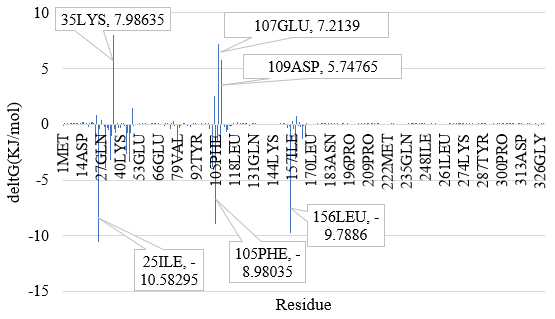


**Figure S3.** 2D representation of hydrogen bonds and hydrophobic interactions. Dashed lines indicate hydrogen bonds, while spiked residues form hydrophobic interactions with CDK9 complexes: a, 1; b: 2; c: 5 (figures were prepared using LIGPLOT).

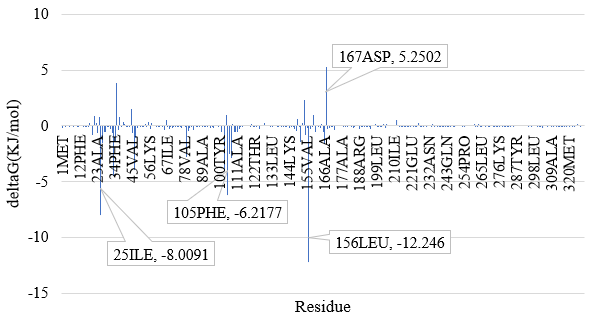


**Figure S4.** Time-dependent RMSD of CDK9/Inhibitor complexes during MD simulation. The plot shows the RMSDs of the three inhibitors relative to their initial minimized structures over the course of MD simulations for the three CDK9/inhibitor complexes.

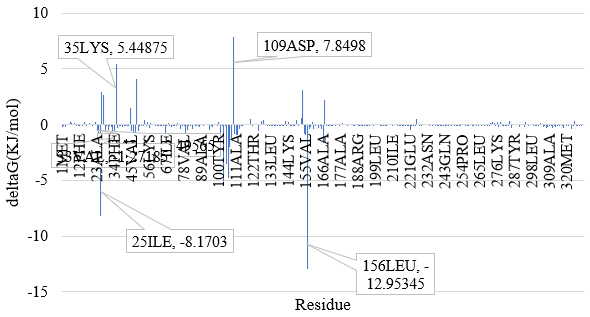
1. CDK9/1



1. CDK9/2



1. CDK9/5



**Figure S5.** Per-Residue decomposition of ∆G for protein−inhibitor complexes. The graphs illustrate the decomposition of binding free energy (∆G) on a per-residue basis for complexes between CDK9 and (a), CDK9/1; (b), CDK9/2; (c), CDK9/5.

**Table S3.** Binding free energy components for the protein–inhibitor complexes calculated using MM-PBSA.a

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| System | ΔGbind | ΔGvdw + ΔGele | ΔGpol | ΔGnonpolar | ΔGsol |
| CDK9/1 | -116.418 | -234.707 | 145.0457 | -26.7564 | 118.2893 |
| CDK9/2 | -109.776 | -213.929 | 129.3147 | -25.1616 | 104.1531 |
| CDK9/5 | -106.895 | -286.115 | 207.6938 | -28.4729 | 179.2209 |

**Table S4.** Hydrogen bonds formed between inhibitors and CDK9 during MD simulations (The latter 20 ns) (Occupancy ≥ 50%)

|  |  |  |  |
| --- | --- | --- | --- |
| System | Donor | Acceptor | Occupancy (%) |
| CDK9/1 | :lig@N4 | :Cys106@H-:Cys106@N | 71.65 |
| :Cys106@O | :lig@H-:lig@N3 | 50.92 |
| CDK9/2 | :lig@N6 | :Cys106@H-:Cys106@N | 52.47 |
| CDK9/5 | :lig@N6 | :Cys106@H-:Cys106@N | 92.2 |
| :Cys106@O | :lig@H-:lig@N4 | 56.07 |