*Nucleic Acids Research*, 2025, **53**, W258–W265
https://doi.org/10.1093/nar/gkaf366
Advance access publication date: 1 May 2025
**Web Server issue**

# Dr. Kinase: predicting the drug-resistance hotspots of protein kinases

**Shaofeng Lin**[1,†], **Chao Tu**[2,3,†], **Ruifeng Hu**[4,†], **Haiji Wang**[5], **Zongcheng Dong**[1], **Hui Luo**[1], **Lan Kuang**[2], **Tao Wang**[2], **Liming Wang**[5], **Zhongming Zhao**[4,*], **Zhihong Li**[2,3,*], **Haodong Xu**[2,3,*]

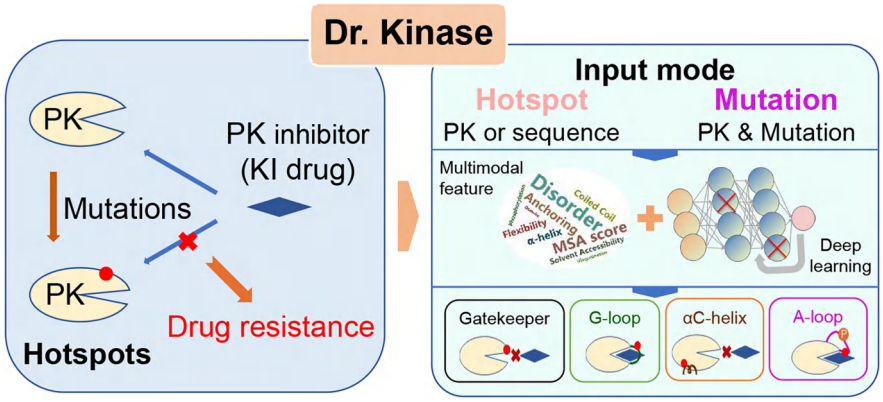
[1]Key Laboratory of Gastrointestinal Cancer (Fujian Medical University), Ministry of Education, School of Basic Medical Sciences, Fuzhou 350122, China
[2]Department of Orthopaedics, The Second Xiangya Hospital, Central South University, Changsha, Hunan 410011, China
[3]Hunan Key Laboratory of Tumor Models and Individualized Medicine, The Second Xiangya Hospital of Central South University, Changsha 410011, China
[4]Center for Precision Health, McWilliams School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX 77030, United States
[5]School of Biomedical Science, Hunan University, Changsha 410082, Hunan, China

[*]To whom correspondence should be addressed. Email: xuhaodong@csu.edu.cn
Correspondence may also be addressed to Zhongming Zhao. Email: zhongming.zhao@uth.tmc.edu
Correspondence may also be addressed to Zhihong Li. Email: lizhihong@csu.edu.cn
[†]The first three authors should be regarded as Joint First Authors.

## Abstract

Protein kinases (PKs) regulate various cellular functions, and are targeted by small-molecule kinase inhibitors (KIs) in cancers and other diseases. However, drug resistance (DR) of KIs occurs through critical mutations in four types of representative hotspots, including gatekeeper, G-loop, αC-helix, and A-loop. KI DR has become a common clinical complication affecting multiple cancers, targeted kinases, and drugs. To tackle this challenge, we report an upgraded web server, namely Dr. Kinase, for predicting the loci of four DR hotspots and assessing effects of mutations on DR hotspots for PKs in our previous studies, by utilizing multimodal features and deep hybrid learning. The performance of Dr. Kinase has been rigorously evaluated using independent testing, demonstrating excellent accuracy with area under the curve values exceeding 0.89 in different types of DR hotspot predictions. We further conducted *in silico* analyses to evaluate and validate the epidermal growth factor receptor mutations on protein conformation and KIs' binding efficacy. Dr. Kinase is freely available at http://modinfor.com/drkinase, with comprehensive annotations and visualizations. We anticipate that Dr. Kinase will be a highly useful service for the basic, translational, and clinical community to unveil the molecular mechanisms of DR and the development of next-generation KIs for emerging cancer precision medicine.

## Graphical abstract




## Introduction

Protein kinases (PKs), as phosphorylation enzymes for modifying proteins, regulate almost all biological processes and cellular functions [1, 2]. Thus, the dysregulation of PKs is highly associated with various human diseases, especially cancer [3]. Therefore, PKs have been prominent targets for modern drug

**Received:** March 5, 2025. **Revised:** April 13, 2025. **Editorial Decision:** April 16, 2025. **Accepted:** April 23, 2025

discovery against many diseases, and these targeted therapies greatly improved patient survival in several cancer: > 400 kinase inhibitors (KIs) that primarily target different activated PKs are currently undergoing clinical trials, and 82 have been approved against many diseases, by the Food and Drug Administration (FDA) as of September 2024 [4, 5]. However, drug resistance (DR) of KIs through critical mutations has become a common clinical complication, affecting multiple cancers, targeted kinases, and drugs.

In the past decade, four types of representative KI-resistant mutation hotspots, including gatekeeper [6], G-loop [7], αC-helix, and A-loop [8], have been identified as actionable and secondary mutations of PKs. Gatekeeper residue is a single amino acid located near the protein–drug binding site to impede drug–PK binding. One notable gatekeeper is the T315I gatekeeper mutation in ABL1, which is resistant to imatinib, dasatinib, and nilotinib [6]. In addition, G-loop, αC-helix, and A-loop occur in the regions for transitions of active and inactive kinase conformations. G-loop is also named glycine-rich loop and phosphorylation loop, and has a conserved consensus motif GxGxxG. E255K/V and Y257C of BCR-ABL are the representative G-loop mutations for disrupting an electrostatic triad required for imatinib binding and causing imatinib resistance [7]. αC-helix, also called C-helix and αC, is a single α-helix located in the N-lobe of the kinase domain between β3 and β4 strands. A-loop, which is called the activation loop or T-loop, is located in the C-lobe and contains a phosphorylation site. Therefore, the systematic investigation of kinase mutations and function on DR hotspots can contribute to molecularly targeted therapies for precision medicine. We developed a kinase mutations and drug response database, KinaseMD, curated four types of PK substructures [9], and shaped the landscape of DR mutations in kinase hotspots [10].

In this study, we present a timely web service, named Dr. Kinase, that predicts the loci of four types of DR hotspots for PKs and quantifies mutation effects for these regions. Dr. Kinase extends our prior findings [9, 10], and utilizes the advantages of deep hybrid learning technology and multimodal features to predict actionable DR hotspots. The performance of Dr. Kinase has been rigorously evaluated using five-fold cross-validation (CV) and independent testing, demonstrating excellent accuracy with area under the curve (AUC) values exceeding 0.89 in different types of DR hotspot predictions. We further performed computational analyses to elucidate the impact of epidermal growth factor receptor (EGFR) and ABL1 mutations in the gatekeeper loci on protein structure and KIs' binding efficiency. Moreover, the web service of Dr. Kinase provides comprehensive annotations and visualizations for the structural and physicochemical features of predicted DR hotspots, and can be accessed freely at http://modinfor.com/drkinase. We anticipate that Dr. Kinase can be helpful for further analysis of DR and PKs.

## Materials and methods

### Data preparation

We obtained and processed the substructure locus information of 547 human kinases from our KinaseMD [9] database. After removing duplicated kinases and filtering out the kinases without specific locus information of the four categories mentioned above, we obtained 388 unique human kinases. Among them, 344, 172, 231, and 312 kinases had 346 gatekeeper, 179 G-loop, 238 αC-helix, and 403 A-loop locus information, respectively. For example, gatekeeper residues (single amino acids like EGFR T790) were identified based on their spatial role in ATP-binding pockets, while dynamic regions, i.e. G-loop/αC-helix/A-loop, always hold conserved lengths (e.g. A-loop: 10–25 residues). The G-loop is structurally defined by a conserved GxGxxG consensus motif, where glycine (G) residues constitute critical anchor points and "x" positions exhibit amino acid variability. For characterization and prediction of cancer DR hotspots, we developed a matched reference dataset as negative through generation of control sequences. Negative samples were designed to avoid functional overlap: Gatekeeper negatives (number: 1720) were randomly selected single residues outside known positions, while G-loop/αC-helix/A-loop negatives (number: 534, 1185, and 2015) were length-matched non-functional segments from unrelated kinase regions. To facilitate the development and evaluation of our models, the constructed dataset was partitioned into a training dataset, which represented 90% of the total data, and an independent dataset, which accounted for the remaining 10% of the data. Furthermore, a kinase-level blind splitting strategy was implemented to assess the model's capacity to generalize unencountered kinases. Test sets contained kinases absent from the training data (e.g. testing on kinase 1 while training on kinase 2, kinase 3, and others, and ensuring the ratio of training and test is still 9:1).

### Model architecture

Dr. Kinase's deep learning framework adopts a bipartite structure combining sequential pattern recognition and feature learning capabilities. The first module employs dynamic sequence representations through SeqVec—a protein-optimized implementation of ELMo's contextual embedding framework. Unlike standard word embeddings, this biological adaptation utilizes bidirectional long short-term memory (LSTM) networks trained on UniRef50's massive unannotated protein sequences to derive position-aware residue encodings. Through unsupervised learning on evolutionary-scale data, the system distills biophysical characteristics and long-range interdependencies into 1024D vectors that holistically represent each amino acid's chemical, structural, and functional attributes. These context-rich embeddings subsequently feed into a parallel feature extraction network. A convolutional neural network branch applies multi-scale filters with max-pooling to detect conserved local patterns indicative of DR hotspot signals, such as sequence motifs. Simultaneously, a bidirectional LSTM processes sequential relationships across variable-length peptide segments, enhanced by trainable position-specific embedding vectors that complement the pre-trained SeqVec features. The architecture implements cross-modal fusion through stacked dense layers that hierarchically integrate spatial-convolutional features with temporal-recurrent patterns. Final classification occurs through a fully connected output layer distinguishing DR hotspots from non-functional peptide fragments. Training employed Adam optimization (learning rate = 0.001, beta_1 = 0.9, and beta_2 = 0.999) with early stop control (patience = 10).

### Model evaluation

The performance of Dr. Kinase was comprehensively evaluated with a rigorous validation framework that encompasses

both CV techniques and independent test sets. First, five-fold CV, a widely used technique in machine learning, assessed the model's performance on the training dataset. Then, receiver operating characteristic (ROC) curves and the corresponding AUC values were constructed for each CV, and five ROC curves were generated, and subsequently, a mean ROC curve was calculated, ensuring that each of the five models carried equal weight. In order to assess the model's generalizability, ROC curves and AUC values were calculated using independent datasets. Additionally, we employed PremPLI [11] to quantify the binding free energy of interaction between the KIs and PKs, thereby characterizing the alterations in binding affinity and conformation induced by actionable DR mutations.

### DR analysis

As previously described [9, 10], the predicted PK hotspots were mapped to mutation loci in the Genomics of Drug Sensitivity in Cancer (GDSC) [12] dataset. Treatment groups were defined as combinations of *Kinase:Drug:Cancer:Sub-tissue*. Each treatment group could contain multiple cell lines, with each line potentially harboring specific mutations. Within each treatment group, cell lines were divided into two comparison groups:

Group 1 (G1): Cell lines with mutations in the identified substructure hotspots, or in the second setting, cell lines that possess secondary mutations in addition to a primary mutation in any of the four substructure regions.

Group 2 (G2): Cell lines without mutations in any of the four substructure regions, or in a second setting, cell lines with a primary mutation in the substructures.

To assess drug resistance, the $IC_{50}$ values were calculated for each group. Increased resistance was inferred when the average $IC_{50}$ of G1 was greater than that of G2 [(i.e. median($IC_{50}$, G1) > median($IC_{50}$, G2) and fold change (FC) $\geq$ 1.5)].

### Functional annotations

The basic information of source protein was obtained to display from the UniProt database [13]. The mutation datasets were curated and mapped from five cancer-related resources, including The Cancer Genome Atlas [14], International Cancer Genome Consortium [15], Catalogue of Somatic Mutations in Cancer [16], Cancer Cell Line Encyclopedia [17], and GDSC [12]. Then, the structure information of proteins is mapped from the Protein Data Bank (PDB) database [18]. The kinase–substrate relations were collected from PhosphoSitePlus [19], to model the kinase–substrate network. Our computational framework quantified 10 biophysical descriptors across the loci of DR hotspots. Structural dynamics profiling incorporated amino acid flexibility profiles derived through DynaMine software [20]. Secondary structure elements (α-helical/coiled-coil) and surface exposure indices were calculated via SPIDER2 [21]. Intrinsic disorder quantification was performed using IUPred algorithms [22]. ANCHOR-based binding site prediction [23] provided interfacial interaction potentials for DR hotspots. Evolutionary conservation metrics originated from orthologous sequence alignments generated through BioWare's Gopher platform [24]. Domain annotations were cross-referenced against Pfam's curated repository [25]. Post-translational modification landscapes were systematically mapped by integrating two resources: phosphorylation events from the Eukaryotic Phosphorylation Sites Database [26] and lysine modification records extracted from Protein Lysine Modification Database (PLMD) [27].

### Web server implementation

The standard model-view-controller (MVC) framework, a prevalent approach in contemporary web application design [28, 29], was used to develop the web server of Dr. Kinase. Three core logical components, namely "Prediction," "Results," and "Controller," collectively constitute the Dr. Kinase system. On the backend, the prediction system controls the execution of submitted jobs to predict DR hotspots of PKs. On the frontend, the "Prediction" module enables user interactions with the system through two prediction modes, "Hotspot mode" and "Mutation mode." Then, the "Controller" component plays a pivotal role in the validation of input data's format, transfer of data from the frontend to the backend interface, and execution of predictive models, and delivering the results to the "Results" page. The "Results" page provides the feature properties and functional annotations for DR hotspots and source PKs. Specifically, the properties of DR hotspots calculated by Dr. Kinase are provided for users. An amalgamation strategy of HTML5, CSS3, JavaScript, and PHP was adopted to ensure a responsive server. jQuery and zTree, JavaScript libraries, were employed to leverage Ajax technology for seamless communication. Additionally, 3Dmol.js [30], ProViz tool [31], and Cytoscape.js [32] were used for the presentation of additional annotations.

## Results

### Overview of the Dr. Kinase framework

Dr. Kinase was designed specifically to predict the locus of four DR hotspots for PKs and infer mutation effects for these regions in our previously published studies [9, 10], utilizing the advantages of deep hybrid learning technology and multimodal features (Fig. 1). First, we curated the substructure locus information of gatekeeper, G-loop, αC-helix, and A-loop from our KinaseMD [9] database (Fig. 1). The peptides centered around the hotspot region were processed as input. Subsequently, multimodal features, integrating sequence, evolution, and structure, were considered to extract for model training (Fig. 1). Then, we utilized a deep hybrid architecture, comprising a protein language model, word embedding, convolution, and BLSTM, to leverage the full potential of these advanced networks and their ability to extract high-level features from protein sequences (Fig. 1). Specifically, the deep learning architecture (e.g. layer type, output shape) is schematically depicted in Supplementary Fig. S1. Furthermore, by having this AI model built in, our tool is able to assess and quantify the effects of mutations on these important functional regions. Last, the interactive web server, comprising the prediction, presentation of results, functional annotations, and visualizations of DR hotspots, was constructed for users (Fig. 1). Taken together, Dr. Kinase provides a comprehensive server for researchers to functional and applied aspects of DR hotspots.

### Performance evaluation of Dr. Kinase

Based on the collected information on drug resistance hotspots, we analyzed the amino acid preference and length of different hotspots (Supplementary Fig. S2). More specif-

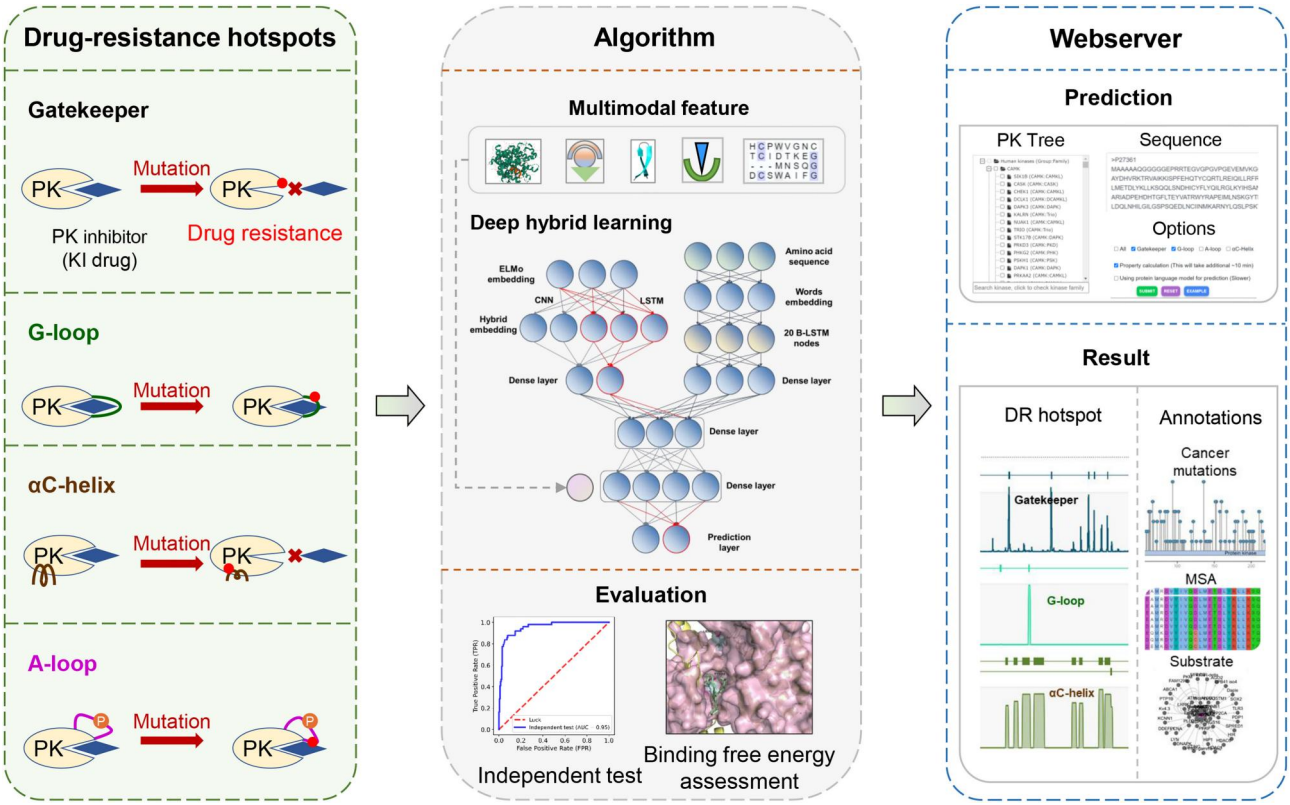



**Figure 1.** The overall framework of Dr. Kinase. Four types of DR hotspots, including gatekeeper, G-loop, αC-helix, and A-loop, were curated. After the encoding of multimodal features, a deep hybrid learning algorithm was designed and evaluated. Then, the web server of Dr. Kinase was constructed for prediction, annotations, and visualizations for DR hotspots.

ically, there are significantly more methionine (M), threonine (T), leucine (L), and phenylalanine (F) residues on gatekeeper (Supplementary Fig. S2A). In addition, we found that L-glutamate (E), valine (V), isoleucine (I), and glycine (G) were significantly enriched upstream and downstream of the gatekeeper (Supplementary Fig. S2B). In addition, for A-loop, G-loop, and αC-helix, we count their most likely sequence lengths (Supplementary Fig. S2C–E). For example, the most representative lengths of A-loop are 24 and 12 (Supplementary Fig. S2C), which is also the length of the peptide library they cut into when making their predictions.

Understanding the structural characteristics of drug resistance hotspots is crucial for predicting their resistance potential and unraveling their functional implications. To characterize the structure characteristics of drug resistance hotspots, we analyzed the curated, experimentally validated hits by comparing them to the background dataset comprising a considerable number of randomly selected peptides with the same length as the drug resistance hotspots. We employed various structural bioinformatics algorithms and tools to identify common structural properties (Fig. 2 and Supplementary Fig. S3). Gatekeeper and its surrounding sequences were more evolutionarily conservative and enriched in functional domains (Fig. 2A). Remarkably, the known gatekeepers exhibited a lower degree of solvent accessibility and protein disorder compared to the random peptides (Fig. 2A). Furthermore, known gatekeepers were found to be preferentially located in higher flexibility regions and had higher binding stability (Fig. 2A). Additionally, the analysis revealed a specific preference of gatekeeper for coiled-coil and sheet regions rather than α-helix regions (Fig. 2A). Sequences of A/G loops were more evolutionarily conservative and enriched in functional domains as well (Fig. 2A). On the contrary, they were found to be preferentially located in protein disordered regions, highlighting their distinctive localization patterns (Fig. 2A). For αC-helix, the analysis revealed a specific preference of α-helix regions rather than coiled-coil and sheet regions (Fig. 2A). It was also observed that αC-helix tends to occur in lower flexibility regions. These findings provided valuable insights into the structural characteristics of drug resistance hotspots and indicate potential determinants for their recognition.

The performance of Dr. Kinase was evaluated by the AUC values through the five-fold CV approach (Supplementary Fig. S4) and demonstrated its great predictive capabilities. Dr. Kinase obtained an average AUC value of 0.96 in predicting four types of DR hotspots, ranging from 0.92 to 0.99 (Supplementary Fig. S3, gatekeeper: 0.99, G-loop: 0.95, A-loop: 0.96, αC-helix: 0.92), indicating consistent and reliable performance. Additionally, Dr. Kinase achieved an average AUC value of 0.93 among different DR hotspots (Fig. 2B, gatekeeper: 0.98, G-loop: 0.95, A-loop: 0.91, αC-helix: 0.89), validated with an independent testing dataset. To further evaluate the model's ability to handle unseen kinases, we performed kinase-level blind splitting, and obtained an average AUC of 0.98 among different DR hotspots (Supplementary Fig. S5, gatekeeper: 0.99, G-loop: 0.94, A-loop: 0.99, αC-helix: 0.98) with an independent testing dataset. These results suggested the robustness and accuracy of Dr. Kinase in predicting DR hotspots in PKs.

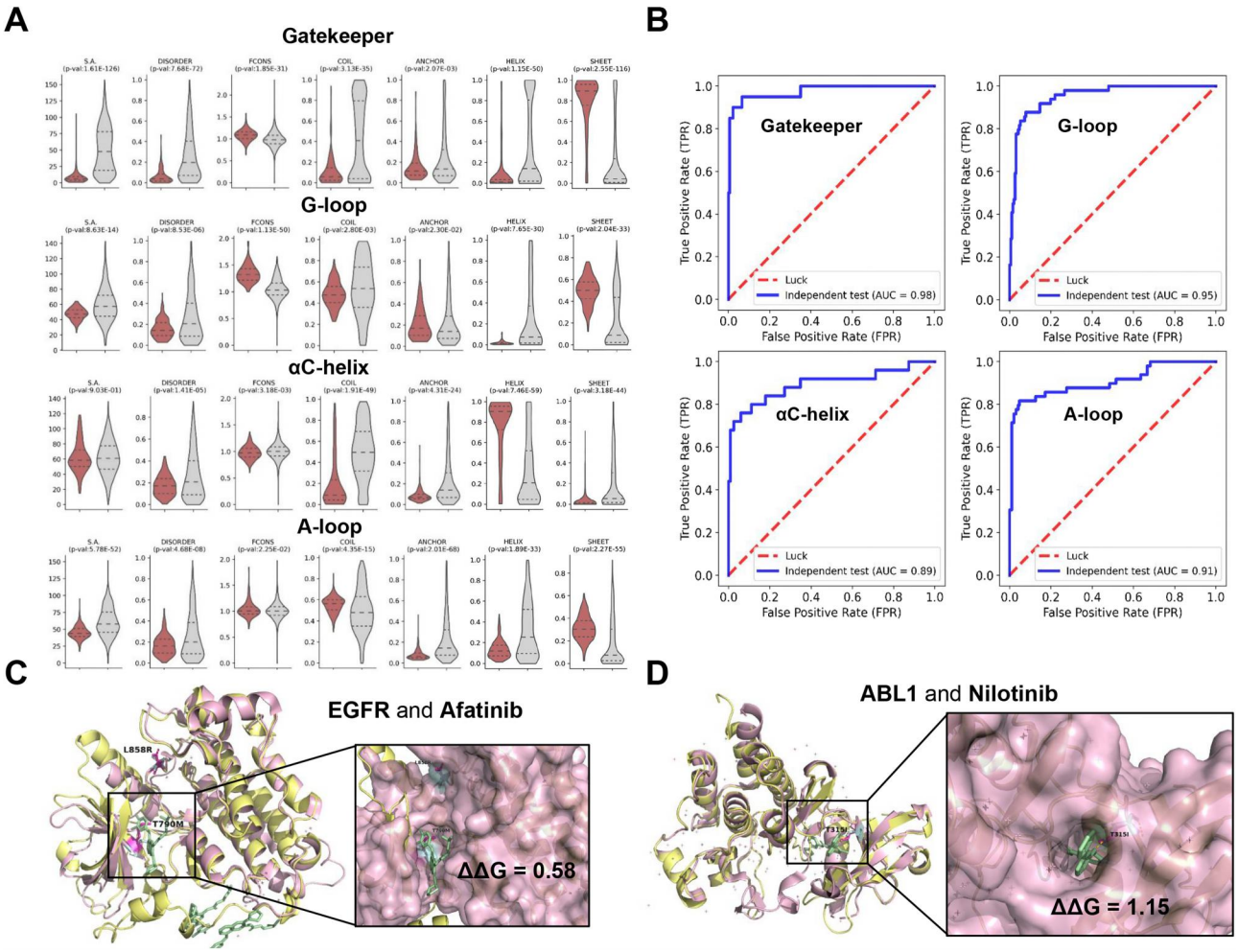



**Figure 2.** Performance of Dr. Kinase based on the deep hybrid learning. (**A**) The statistics of multiple characteristic features, including solvent accessibility (S.A.), disorder, flanking conservation (FCONS), coil, anchor, α-helix, and β-sheet, for the known four types of DR hotspot instances and random peptides. (**B**) ROC curves of Dr. Kinase in an independent testing dataset. Binding free energy analysis for the drug afatinib (PDB code: 4G5J) that binds to EGFR (**C**) and nilotinib (PDB code: 3CS9) that binds to ABL1 (**D**).

To delineate mutation-induced alterations in kinase-inhibitor interactions, we performed binding free energy calculations using PremPLI [11] on wild-type versus mutant PKs. For EGFR-targeting agents, afatinib (PDB: 4G5J), forming covalent bonds with CYS797, exhibited reduced binding affinity ($\Delta\Delta G$ = +0.58 kcal/mol) under L858R/T790M double mutation (Fig. 2C). Gefitinib (PDB: 2ITY) exhibited an analogous destabilization result (Supplementary Fig. S6), confirming mutation-driven resistance mechanisms across EGFR inhibitors. In ABL1 complexes, nilotinib (PDB: 3CS9) demonstrated affinity loss ($\Delta\Delta G$ = +1.15 kcal/mol) upon T315I mutation (Fig. 2D), consistent with clinical resistance profiles. Both of the above gatekeeper loci can be predicted by Dr. Kinase (they are excluded from the training set).

## The usage of Dr. Kinase web server

The web server of Dr. Kinase was designed and developed specifically for predicting four DR hotspots of PKs, with a modular and user-friendly manner (Fig. 3 and Supplementary Fig. S7). Two prediction modes, "Hotspot mode" and "Mutation mode," were alternative to predict the loci of four DR hotspots and assess the effects of mutations on DR hotspots, respectively. For the former, two formats, selection from a clickable and searchable hierarchical classification tree of PKs and input with PK sequence(s), are used to input single or multiple PKs (Fig. 3A). Then, one or more hotspot types can be submitted, more parameters are optional for the calculation of property or choice of protein language model (Fig. 3A). The prediction results will be automatically visualized by a 3D structural model, a line chart, and a tabular list with detailed information, including "DR hotspot type," "Hotspot instance," "Position," "Score," and "Detail" (Fig. 3B). And Then, the structural and physicochemical features of predicted DR loci and information of source protein (Fig. 3C and D) could be viewed by clicking the "Detail" button.

In addition to the basic predictions, integrated annotations of DR hotspots and source PKs, including cancer mutations, 3D structure, multiple sequence alignment (MSA) in species, and kinase–substrate networks, were processed and visualized (Fig. 3D–G). The structure of DR hotspot and source protein is presented with 3Dmol.js [30] (Fig. 3D), and the cancer mutation map was represented with a needle plot (Fig. 3E). Moreover, the MSA of DR hotspot and source protein are visualized to view the sequence conservation with the ProViz tool [31] (Fig. 3F), and the substrates of source PK are provided in a table and an interactive network with the Cytoscape.js [32] (Fig. 3G). Moreover, the "Mutation mode" was


**Figure 3.** The usage of the Dr. Kinase web server. (**A**) The example input of PKs and DR hotspot types in the "Prediction" page. (**B**) The prediction results for the example. (**C**) The structural and physicochemical features of the DR hotspot. The annotations for predicted DR hotspots and source PK, including basic information and structural visualization (**D**), cancer mutation map (**E**), MSA viewer (**F**), and the substrate network (**G**).

constructed to infer mutation effects for four DR hotspots, based on the PK selection from the clickable and searchable hierarchical classification tree and mutation input of selected PKs (Supplementary Fig. S7A). After submitting jobs, the score changes before and after mutation, as well as 3D structural model and line chart that were returned to further infer actionable mutations may be associated with DR (Supplementary Fig. S7B). Taken together, Dr. Kinase is a useful server for predicting four DR hotspots of PKs, allowing researchers to study DR mechanism of KIs and develop next-generation KIs.

## Discussion

PKs are involved in almost all biological processes and have been prominent targets for modern drug discovery against cancer and diseases. However, PKs targeted by small-molecule inhibitors develop resistance through mutations in four types of representative mutation hotspots (gatekeeper, G-loop, αC-helix, and A-loop). In this regard, the identification of DR hotspots is the foundation for understanding the mechanisms of DR and developing next-generation KIs. In this study, we report a web server, called Dr. Kinase, for predicting four DR hotspots for PKs, through the integrated multimodal features and deep hybrid learning strategy. Notably, Dr. Kinase showed excellent accuracy for different types of DR hotspot predictions, in both five-fold CV and independent testing. Furthermore, Dr. Kinase web server also provides functional annotations and visualizations of structural and physicochemical features for the DR hotspots and source proteins. In addition, for any mutation within a potential DR-related region, the model calculates the score change before and after mutation, and further infers actionable mutations may be associated with DR. In summary, Dr. Kinase would serve users to predict DR hotspots on PKs, for further analysis of DR mechanism and drug development.

We further demonstrated the tool application by integrating our tool's prediction with the GDSC dataset (drug response data for 621 compounds across 1000+ cancer cell lines), which revealed new drug resistance events linked to mutations in predicted hotspots. We first used a model to predict DR hotspots across the entire human kinome. This revealed potential functional regions with drug resistance potential (Supplementary Table S1). By integrating these predictions with the drug sensitivity data in the GDSC, we identified over 1000 drug resistance events, i.e. increased resistance (FC $\geq$ 1.5), for treatment groups linked to mutations in predicted hotspots (Supplementary Table S1). For example, EGFR L858R/T790M double mutation showed a $\Delta\Delta G$ = +0.58 kcal/mol for afatinib and gefitinib binding, aligning

with GDSC-derived $IC_{50}$ shifts (>20-fold and 800-fold resistance). The pan-kinome predictions also enable pre-screening patient mutations to avoid therapies prone to resistance.

In the future, we will continuously extend the benchmark dataset and improve the performance of Dr. Kinase when newly identified DR hotspots are discovered. Meanwhile, we will adopt more useful features and machine learning frameworks to improve Dr. Kinase models. Moreover, the web server of Dr. Kinase will be persistently maintained and improved, such as interface and visualization.

## Acknowledgements

*Author contributions*: Shaofeng Lin (Methodology, Visualization, Writing—original draft), Chao Tu (Data curation, Investigation, Software), Ruifeng Hu (Formal analysis, Software, Visualization) Haiji Wang, Zongcheng Dong, Hui Luo (Validation, Visualization) Lan Kuang, Tao Wang, Liming Wang (Visualization), Zhongming Zhao, Zhihong Li, and Haodong Xu (Conceptualization, Supervision, Writing—review & editing)

## Supplementary data

Supplementary data is available at NAR online.

## Conflict of interest

None declared.

## Funding

This work was supported by the Natural Science Foundation of China [32300528, 32300520, and 82272664]; the Science and Technology Innovation Program of Hunan province [2023RC3080 and 2023RC3085]; the Excellent Youth Foundation of Hunan Scientific Committee [2024JJ2084]; the Scientific Research Fund of Hunan Provincial Education Department [23B0023 and 24A0008]; Hunan Provincial Health High-Level Talent Scientific Research Project (R2023054); and Scientific Research Foundation for Advanced Talents of Fujian Medical University [XRCZX2022015]. Funding to pay the Open Access publication charges for this article was provided by the Natural Science Foundation of China [32300528].

## Data availability

Dr. Kinase can be accessed freely at http://modinfor.com/drkinase and the source at https://github.com/BioDataStudy/DrKinase (permanent DOI: 10.5281/zenodo.15200413).

## References

1. Manning G, Whyte DB, Martinez R *et al.* The protein kinase complement of the human genome. *Science* 2002;**298**:1912–34. https://doi.org/10.1126/science.1075762
2. Chen M, Zhang W, Gou Y *et al.* GPS 6.0: an updated server for prediction of kinase-specific phosphorylation sites in proteins. *Nucleic Acids Res* 2023;**51**:W243–50. https://doi.org/10.1093/nar/gkad383
3. Manning BD, Toker A. AKT/PKB signaling: navigating the network. *Cell* 2017;**169**:381–405. https://doi.org/10.1016/j.cell.2017.04.001
4. Roskoski R Jr. Properties of FDA-approved small molecule protein kinase inhibitors: a 2024 update. *Pharmacol Res* 2024;**200**:107059. https://doi.org/10.1016/j.phrs.2024.107059.
5. Carles F, Bourg S, Meyer C *et al.* PKIDB: a curated, annotated and updated database of protein kinase inhibitors in clinical trials. *Molecules* 2018;**23**:908. https://doi.org/10.3390/molecules23040908
6. Miller GD, Bruno BJ, Lim CS. Resistant mutations in CML and Ph(+)ALL—role of ponatinib. *Biologics* 2014;**8**:243–54. https://doi.org/10.2147/btt.S50734
7. Shah NP, Nicoll JM, Nagar B *et al.* Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* 2002;**2**:117–25. https://doi.org/10.1016/s1535-6108 (02)00096-x
8. Scheeff ED, Eswaran J, Bunkoczi G *et al.* Structure of the pseudokinase VRK3 reveals a degraded catalytic site, a highly conserved kinase fold, and a putative regulatory binding site. *Structure* 2009;**17**:128–38. https://doi.org/10.1016/j.str.2008.10.018
9. Hu R, Xu H, Jia P *et al.* KinaseMD: kinase mutations and drug response database. *Nucleic Acids Res* 2021;**49**:D552–61. https://doi.org/10.1093/nar/gkaa945
10. Kim P, Li H, Wang J *et al.* Landscape of drug-resistance mutations in kinase regulatory hotspots. *Brief Bioinform* 2021;**22**:bbaa108. https://doi.org/10.1093/bib/bbaa108
11. Sun T, Chen Y, Wen Y *et al.* PremPLI: a machine learning model for predicting the effects of missense mutations on protein–ligand interactions. *Commun Biol* 2021;**4**:1311. https://doi.org/10.1038/s42003-021-02826-3.
12. Yang W, Soares J, Greninger P *et al.* Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* 2013;**41**:D955–61. https://doi.org/10.1093/nar/gks1111
13. UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2025. *Nucleic Acids Res* 2025;**53**:D609–17. https://doi.org/10.1093/nar/gkae1010
14. Sanchez-Vega F, Mina M, Armenia J *et al.* Oncogenic signaling pathways in The Cancer Genome Atlas. *Cell* 2018;**173**:321–37. https://doi.org/10.1016/j.cell.2018.03.035
15. Hudson TJ, Anderson W, Artez A *et al.* International network of cancer genome projects. *Nature* 2010;**464**:993–8. https://doi.org/10.1038/nature08987
16. Sondka Z, Dhir NB, Carvalho-Silva D *et al.* COSMIC: a curated database of somatic variants and clinical data for cancer. *Nucleic Acids Res* 2024;**52**:D1210–7. https://doi.org/10.1093/nar/gkad986
17. Ghandi M, Huang FW, Jané-Valbuena J *et al.* Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* 2019;**569**:503–8. https://doi.org/10.1038/s41586-019-1186-3
18. Burley SK, Bhatt R, Bhikadiya C *et al.* Updated resources for exploring experimentally-determined PDB structures and computed structure models at the RCSB Protein Data Bank. *Nucleic Acids Res* 2025;**53**:D564–74. https://doi.org/10.1093/nar/gkae1091
19. Hornbeck PV, Zhang B, Murray B *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res* 2015;**43**:D512–20. https://doi.org/10.1093/nar/gku1267
20. Cilia E, Pancsa R, Tompa P *et al.* The DynaMine webserver: predicting protein dynamics from sequence. *Nucleic Acids Res* 2014;**42**:W264–70. https://doi.org/10.1093/nar/gku270
21. Yang Y, Heffernan R, Paliwal K *et al.* SPIDER2: a package to predict secondary structure, accessible surface area, and main-chain torsional angles by deep neural networks. *Methods Mol Biol* 2017;**1484**:55–63. https://doi.org/10.1007/978-1-4939-6406-2_6

22. Erdős G, Pajkos M, Dosztányi Z. IUPred3: prediction of protein disorder enhanced with unambiguous experimental annotation and visualization of evolutionary conservation. *Nucleic Acids Res* 2021;**49**:W297–303. https://doi.org/10.1093/nar/gkab408
23. Mészáros B, Simon I, Dosztányi Z. Prediction of protein binding regions in disordered proteins. *PLoS Comput Biol* 2009;**5**:e1000376. https://doi.org/10.1371/journal.pcbi.1000376.
24. Lee TJ, Pouliot Y, Wagner V *et al.* BioWarehouse: a bioinformatics database warehouse toolkit. *BMC Bioinformatics* 2006;**7**:170. https://doi.org/10.1186/1471-2105-7-170
25. Mistry J, Chuguransky S, Williams L *et al.* Pfam: the protein families database in 2021. *Nucleic Acids Res* 2021;**49**:D412–9. https://doi.org/10.1093/nar/gkaa913
26. Lin S, Wang C, Zhou J *et al.* EPSD: a well-annotated data resource of protein phosphorylation sites in eukaryotes. *Brief Bioinform* 2021;**22**:298–307. https://doi.org/10.1093/bib/bbz169
27. Xu H, Zhou J, Lin S *et al.* PLMD: an updated data resource of protein lysine modifications. *J Genet Genomics* 2017;**44**:243–50. https://doi.org/10.1016/j.jgg.2017.03.007
28. Xu H, Hu R, Jia P *et al.* 6mA-Finder: a novel online tool for predicting DNA $N^6$-methyladenine sites in genomes. *Bioinformatics* 2020;**36**:3257–9. https://doi.org/10.1093/bioinformatics/btaa113
29. Zheng M, Lin S, Chen K *et al.* MetaDegron: multimodal feature-integrated protein language model for predicting E3 ligase targeted degrons. *Brief Bioinform* 2024;**25**:bbae519. https://doi.org/10.1093/bib/bbae519
30. Rego N, Koes D. 3Dmol.js: molecular visualization with WebGL. *Bioinformatics* 2015;**31**:1322–4. https://doi.org/10.1093/bioinformatics/btu829
31. Jehl P, Manguy J, Shields DC *et al.* ProViz—a web-based visualization tool to investigate the functional and evolutionary features of protein sequences. *Nucleic Acids Res* 2016;**44**:W11–5. https://doi.org/10.1093/nar/gkw265
32. Franz M, Lopes CT, Fong D *et al.* Cytoscape.js 2023 update: a graph theory library for visualization and analysis. *Bioinformatics* 2023;**39**:btad031. https://doi.org/10.1093/bioinformatics/btad031.



**Received:** March 5, 2025. **Revised:** April 13, 2025. **Editorial Decision:** April 16, 2025. **Accepted:** April 23, 2025