JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/jacsau
Article

Accurate Characterization of Binding Kinetics and Allosteric Mechanisms for the HSP90 Chaperone Inhibitors Using AI-Augmented Integrative Biophysical Studies

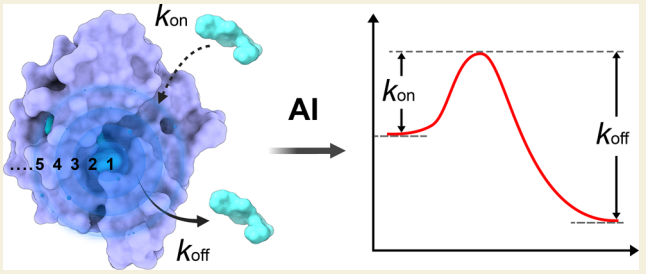
Chao Xu,[#] Xianglei Zhang,[#] Lianghao Zhao, Gennady M. Verkhivker, and Fang Bai*

Cite This: *JACS Au* 2024, 4, 1632−1645

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The binding kinetics of drugs to their targets are gradually being recognized as a crucial indicator of the efficacy of drugs *in vivo*, leading to the development of various computational methods for predicting the binding kinetics in recent years. However, compared with the prediction of binding affinity, the underlying structure and dynamic determinants of binding kinetics are more complicated. Efficient and accurate methods for predicting binding kinetics are still lacking. In this study, quantitative structure−kinetics relationship (QSKR) models were developed using 132 inhibitors targeting the ATP binding domain of heat shock protein 90α (HSP90α) to predict the dissociation rate constant (k_{off}), enabling a direct assessment of the drug−target residence time. These models demonstrated good predictive performance, where hydrophobic and hydrogen bond interactions significantly influence the k_{off} prediction. In subsequent applications, our models were used to assist in the discovery of new inhibitors for the N-terminal domain of HSP90α (N-HSP90α), demonstrating predictive capabilities on an experimental validation set with a new scaffold. In X-ray crystallography experiments, the loop-middle conformation of *apo* N-HSP90α was observed for the first time (previously, the loop-middle conformation had only been observed in *holo*-N-HSP90α structures). Interestingly, we observed different conformations of *apo* N-HSP90α simultaneously in an asymmetric unit, which was also observed in a *holo*-N-HSP90α structure, suggesting an equilibrium of conformations between different states in solution, which could be one of the determinants affecting the binding kinetics of the ligand. Different ligands can undergo conformational selection or alter the equilibrium of conformations, inducing conformational rearrangements and resulting in different effects on binding kinetics. We then used molecular dynamics simulations to describe conformational changes of *apo* N-HSP90α in different conformational states. In summary, the study of the binding kinetics and molecular mechanisms of N-HSP90α provides valuable information for the development of more targeted therapeutic approaches.



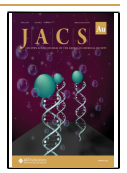
KEYWORDS: *machine learning, drug−target binding kinetics, quantitative structure−kinetics relationship (QSKR), heat shock protein 90 (HSP90), molecular dynamics simulations*

INTRODUCTION

Improving binding affinity has always been a goal in the drug development process. However, thermodynamic parameters measured in closed systems often fail to precisely reflect the actual effect of a drug *in vivo*.[1,2] Recent studies emphasize the crucial role of drug−target binding kinetics in determining efficacy and safety,[3−6] with parameters such as the association rate constant (k_{on}) and dissociation rate constant (k_{off}) being of fundamental importance. The life cycle of a ligand−receptor complex, expressed as the residence time ($\tau = 0.693/k_{off}$), significantly influences the efficacy and duration of biological action of a drug. This factor is crucial for the clinical success of drug candidates.[4,7−10] For example, the drug tiotropium with a long-residence time has shown a sustained bronchodilator effect and improved receptor selectivity.[11] Conversely, analogues such as surgical procedures and analgesics like sufentanil with a short residence time can increase safety during surgical procedures.[12] Predicting and optimizing the binding kinetics of drug candidates in early phases of drug discovery would be beneficial.

Using artificial intelligence to predict binding kinetics parameters is more challenging than predicting binding affinity due to the limited data availability and the complexity involved in the association and dissociation processes.[13−16] Recently reported methodologies, including molecular dynamics simulations, enhanced sampling techniques, quantitative structure−kinetics relationship (QSKR) models, and machine learning

Received: February 9, 2024
Revised: March 15, 2024
Accepted: March 18, 2024
Published: April 1, 2024



https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645

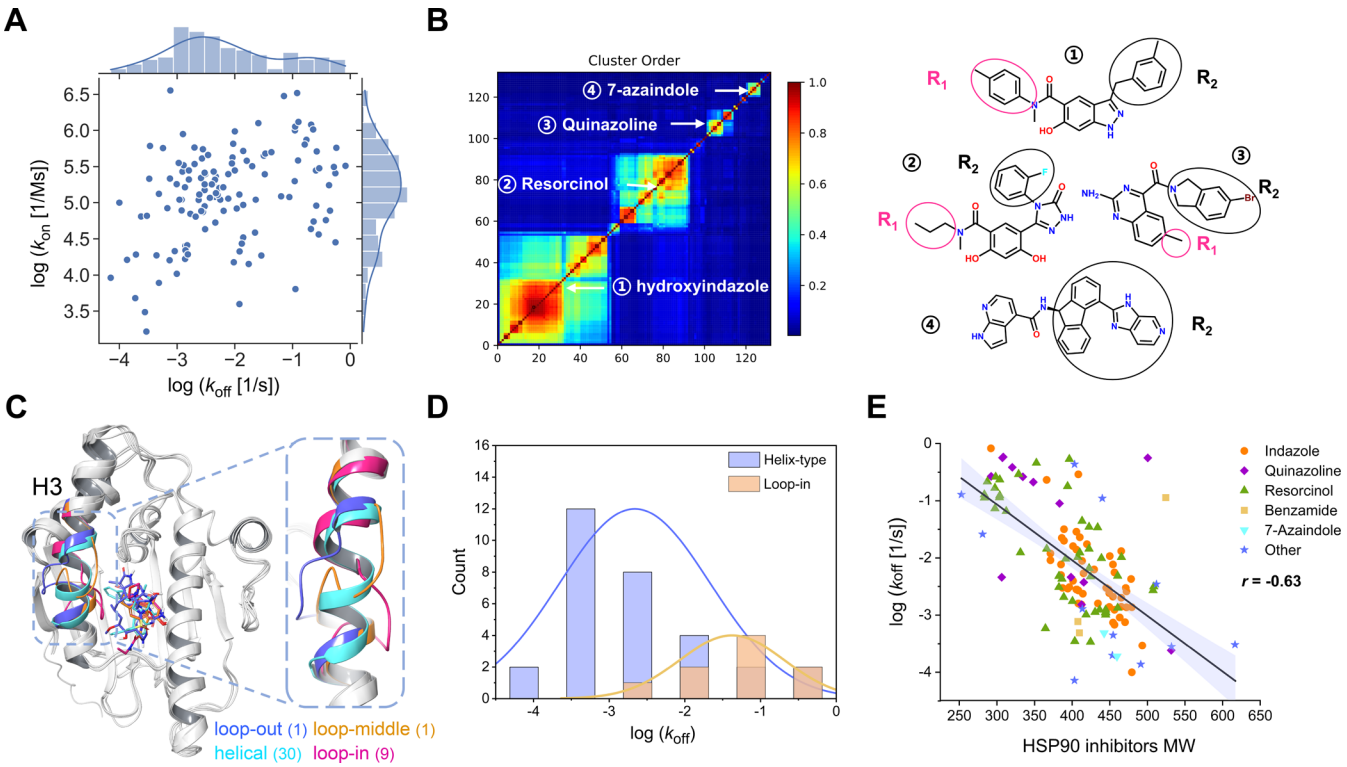


Figure 1. Structure and binding kinetic properties of the 132 N-HSP90α inhibitors. (A) Distribution of binding kinetic parameters for the N-HSP90α inhibitor data set. (B) Similarity matrix of the 132 N-HSP90α inhibitors generated by the Canvas Similarity and Clustering module of Maestro (Schrödinger, 2023). The figure shows the four major scaffold types of compounds, ① hydroxy-indazole, ② resorcinol, ③ quinazoline, and ④ 7-azazindole derivative; The R1 substituent is located near the hydrophilic ATP pocket, and the R2 substituent points to the hydrophobic pocket. (C) The alignment of complex structures of N-HSP90α featuring four different conformations in the data set (PDB IDs: 1OSF, blue; 5J20, cyan; 5NYH, magenta; 3TUH, orange) (the number of compounds is in parentheses). (D) Statistical distribution of dissociation rate constant of inhibitors with loop-in conformation and helical conformation. (E) Correlation analysis between the molecular weight of various inhibitors and the dissociation rate constant.

(ML), have been used for this purpose.[17] Molecular dynamics simulations provide a detailed analysis of the dynamic process of ligand association and dissociation at the atomic level but face a challenge; residence times are usually orders of magnitude longer than the feasible simulation times, greatly limited by their application. Strategies such as metadynamics simulations,[16,18−20] weighted ensemble methods,[21−24] Markov state models,[16,25,26] scaled-MD simulations,[27−29] and random acceleration molecular dynamics (RAMD)[30,31] have emerged to expedite the process. Despite their effectiveness, these methods require significant computational resources, which complicates their application in lead compound discovery. In addition, the proper weighting factors, choice of collective variables, and inherent limitations of the force field and water models can affect simulation accuracy.[32,33] The approach of Bai et al. calculates the association and dissociation free energies by constructing an energy landscape, allowing the simultaneous prediction of binding affinity and kinetic parameters, even in the absence of experimental structural data.[34] The method has recently been applied to distinguish the differences in the binding kinetics of different types of inhibitors of acetylcholinesterase (AChE).[35] However, its applicability is limited to systems without significant conformational changes and requires extensive sampling for accurate prediction. Recently, various molecular modeling techniques for predicting binding affinity have been used for the task of predicting binding kinetics between drugs and targets.[36−40] Ganotra et al. used the COMparative BINding Energy (COMBINE) method and calculated interaction energies for ligand−receptor complexes in data sets that contain 70 N-HSP90α inhibitors and 36 HIV-1 inhibitors. The QSKR model constructed using partial least-squares exhibits good predictive performance, with correlation coefficient of 0.69 and 0.70, and the mean absolute errors (MAEs) were 0.45 and 0.58 on two test data sets, respectively.[38] In two recent studies, predictive models for dissociation rate constant were developed on large data sets including 85 and 155 different protein types, respectively.[41,42] These models were built using random forests with features based on 501 or 680 protein−ligand complexes and protein−ligand atom pairs. The models demonstrated moderate accuracy on the training set (R^2 = 0.6, MAE = 0.59). However, the predictive performance on the test set was lower (R^2 = −0.76, MAE = 1.34 for 100 N-HSP90α inhibitors), potentially attributed to the heterogeneity in experimental data conditions and methods or the complexity of binding and dissociation processes across different protein types and ligands. In summary, the prediction and methodological study of drug−target binding kinetics is an important research focus in computer-aided drug design. Compared to other methods, the QSKR model offers a computationally efficient option for large-scale screening that significantly reduces computational costs. However, due to the limited data availability and the complexity of binding and dissociation processes across different protein types, the practical predictive and applicative capabilities of current methods remain unclear. Efficient and accurate binding kinetics prediction methods await further advancement.

https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645

Heat shock protein 90 (HSP90) is an attractive target for the prediction of binding kinetics properties and has already reported certain numbers of binding kinetics data. This provides an opportunity to explore the application of artificial intelligence algorithms in the task of binding kinetics prediction for ligands against this target. As an ATP-dependent molecular chaperone, HSP90 plays a critical role in fundamental cellular processes and regulatory pathways, making it an attractive target for cancer therapy.[43−48] The ATP binding site in the N-terminal domain of HSP90α (N-HSP90α) is a primary focus for drug development, as the inhibition of HSP90's ATPase activity leads to the degradation of client proteins.[49] The ATP pocket comprises a highly flexible lid region (residues 102−137) that undergoes remarkable conformational changes during the functional cycle of the chaperone, characterized by the opening and closing of the lid. Importantly, this conformational change is not exclusively determined by client protein and nucleotide binding.[50,51] Using solution nuclear magnetic resonance spectroscopy, Henot et al. identified the closed conformation of *apo* N-HSP90α.[52] The residues 104−114 of the ATP-lid are highly plastic and exhibit different "loop-out" and "loop-in" conformations in the absence of a bound ligand.[53] When a ligand is bound, new "loop-middle" and "helical" conformations are formed. Inhibitors not only competitively bind to ATP but also induce local changes in N-HSP90α, which are crucial for the chaperone functional cycle.

As briefly illustrated in Figure S1, this study used machine learning methods to construct a quantitative structure−kinetics relationship (QSKR) model for the dissociation rate constants of HSP90 inhibitors. The QSKR model demonstrated excellent predictive performance (R^2 = 0.93, MAE = 0.18, determination coefficient and mean absolute error) on the test set. To further evaluate the practical predictive ability of the model and its application in the discovery of new N-HSP90α inhibitors, we designed a hit molecule identification strategy by combining the binding affinity-based virtual screening method and the binding kinetics-based filtering module. This combined approach successfully identified 16 previously unreported N-HSP90α inhibitors. Validation and application of these methods underscored their efficacy in predicting binding kinetics. Furthermore, using X-ray techniques and complemented by molecular dynamics simulations, we elucidated the landscape of conformational changes of *apo* N-HSP90α in different states and revealed the relationships between the conformational shifts and the binding kinetics triggered by ligand binding. In summary, a more detailed understanding of the binding kinetics and the interactions between N-HSP90α and inhibitors paves the way for the development of more targeted therapeutic approaches.

MATERIALS AND METHODS

Processing and Analysis of the Data Set

There are 142 N-HSP90α ATP inhibitors collected on the KBbox[54] web server (https://kbbox.h-its.org/). The information includes the structures of the compounds and the binding kinetics values determined by surface plasmon resonance (SPR) measurements as part of the K4DD[8] project. After excluding data with 10 replicates and outliers (data where k_{on} and k_{off} are precisely the same or exceed the sensitivity threshold of SPR measurements) from the data set, the remaining 132 exhibit a broad distribution of binding kinetics spanning different orders of magnitude (Figure 1A). These inhibitors have different scaffold types, and cluster analysis reveals that hydroxy-indazole (55), resorcinol (48), amino-quinazoline (13), and 7-azazindole (2) (the number of compounds is in parentheses) are the most common scaffold types among them. The data set includes crystal structures of a total of 41 inhibitors determined through X-ray analysis. In these structures, the R1 substituents of the different scaffolds are directed toward the entrance of the hydrophilic pocket, while the R2 substituents point to the internal hydrophobic pocket (Figure 1B). N-HSP90α has a relatively flexible ATP-binding pocket. Henot et al. clustered over 300 atomic-resolution structures of N-HSP90α and classified them into eight classes based on a threshold.[52] In this study, we focused on significant conformational differences in the region of residues 104−114, as it profoundly influences the volume and solvent distribution of the ATP-binding pocket, impacting the binding kinetics behavior of ligands.

The binding kinetics properties of different ligands are closely related to the conformation of the complexes. Among the 41 complex structures, residues 104−114 of the ATP-lid display four different conformations (Figure 1C), including different "loop" conformations ("loop-out", "loop-in", and "loop-middle") as well as "helical" conformation. In comparison to inhibitors with a "loop-in" conformation, those inducing a "helical" conformation display slower binding and dissociation rates (Figures 1D and S2A). This may be related to the instability of the transition state and the stabilization of the ground state during the binding process, and the impact of the transition state on residence time has been highlighted in several recent studies.[15,53,55,56] The dissociation rate constant of inhibitors with a loop conformation shows a strong correlation with affinity (r = 0.92) (Figure S2B,C), which may be related to the stabilization of the ground state after binding. Additionally, consistent with previous observations by Kokh et al.,[31] there is a relatively strong correlation (r = 0.63) between the molecular weight of inhibitors in the data set and the dissociation rate constant (Figure 1E) and no obvious correlation observed with the association rate constant (r = 0.44) (Figure S2D). The correlation coefficients for resorcinol and hydroxy-indazole inhibitors with molecular weight and dissociation rate constant are 0.67 and 0.7, respectively. This can be succinctly understood as compounds with slow binding and dissociation rates often feature large hydrophobic moieties. During binding and dissociation, these compounds induce more complex conformational changes in the ATP-lid and can occupy the transient hydrophobic cavity formed between α-helix 3 and β-strand in the ATP pocket, engaging in strong van der Waals interactions mediated by hydrophobic residues.

Preparation of Complex Structures

Complex structures of ligand−receptor are crucial for understanding interaction characteristics, active conformations, dissociation features, and other essential information. To comprehensively characterize and assess the relationship between structure and binding kinetics, we employed a docking approach to obtain complex structures for the remaining 91 inhibitors binding to N-HSP90α in the data set. Due to the flexibility of the ATP-lid in N-HSP90α, particularly at residues 104−114, different inhibitors can give rise to different conformations, posing a major challenge for obtaining rational and accurate complex structures by docking.

In this study, multiple structures of N-HSP90α with four different conformations ("loop-out", "loop-in", "loop-middle", and "helical" conformations) were used as templates for docking with each of the 91 inhibitors. Figure S3 shows the diagram of the docking workflow, specifying the PDB structures used. All molecular docking processes were performed using the Glide module (Schrödinger, 2023). In the graphical interface software Maestro by Schrödinger, the LigPrep module was initially used to process the structures of 91 inhibitors, generating different protonation states, stereochemistry, and tautomers under conditions of pH 7.0 ± 2.0. Subsequently, the Protein Preparation Workflow module was used to optimize the protein structures, including determining the protonation states at pH 7.4 and performing energy minimization using the OPLS4 force field. Finally, the Receptor Grid Generation module was used to generate docking boxes at the ATP-binding pocket, and docking was performed using the Ligand Docking module. In the Glide module (Schrödinger, 2023), a series of hierarchical filters were used to score the binding poses of ligands and receptors.[57] The initial filters tested the spatial fit of the

https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645

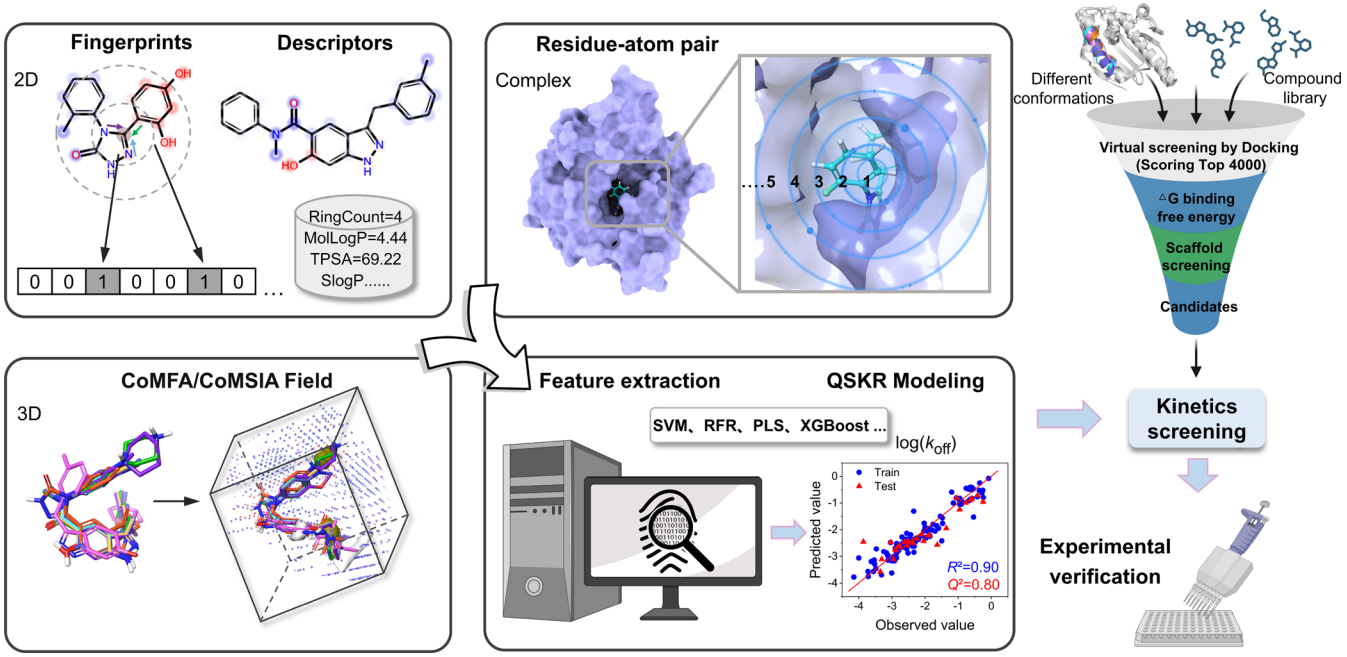


Figure 2. Workflow of QSKR model construction and lead compound screening. (1) Feature extraction methods: features are extracted from molecular SMILES,[67] 3D structures and complex structures. (2) Different machine learning methods are employed to evaluate and construct regression models for k_{off}. (The AutoQSAR model's performance is illustrated in the figure, with $R^2 = 0.90$ on the training set and $R^2 = 0.80$ on the test set.) (3) After docking with different N-HSP90α conformations, further screening is conducted using the QSKR model for predictive selection. The selected compounds are then subjected to surface plasmon resonance (SPR) experiments to determine the dissociation rate constant (k_{off}).

ligand to the defined active site and examined the complementarity of ligand−receptor interactions using a grid-based method. The final scoring defaulted to using Schrödinger's proprietary GlideScore multiligand scoring function. For each conformation, the complex structures obtained by docking were subjected to the Molecular Mechanics Generalized Born Surface Area (MM-GBSA) scoring function to calculate the binding free energy (Schrödinger, 2023). MM-GBSA, a free-energy prediction method that strikes a balance between accuracy and speed, is widely used for affinity-based virtual screening and optimization of docking structures.[58−61] For each inhibitor, the conformation with the lowest binding free energy was selected. We also investigated the similarity between the structures of the 91 inhibitors and the 41 inhibitors with complex structures determined by X-ray. Following the theory that inhibitors with structural similarity binding to the same target have similar binding modes, the binding conformations of the partially similar inhibitors were adjusted and aligned. Ultimately, complex structures of the 91 inhibitors binding to N-HSP90α were obtained. For the 41 complex structures in the data set, we used the Protein Preparation module to determine the protonation states of protein residues and ligands under pH 7.4 conditions, optimize hydrogen bond networks, and perform energy minimization. In the end, 132 complex structures were obtained for feature extraction.

Feature Extraction and Construction of the QSKR Models

In this study, features were extracted from molecular two-dimensional (2D) structures, three-dimensional (3D) structures, and complex structures. The process of feature extraction and the construction and application of QSKR models are illustrated in Figure 2. All machine learning processes were executed using the scikit-learn Python library.[62] Additionally, we used the Maestro software for structure preparation and building the QSKR models. Maestro is Schrödinger's graphical interface that provides access to the state-of-the-art predictive computational modeling and machine learning workflows for molecular discovery.

Feature Extraction Based on Molecular 2D Structures. Initially, we employed the AutoQSAR module in the Maestro software (Schrödinger, 2023) for model construction. This module is an automated tool for quantitative structure−activity relationship modeling, widely applied to predict properties such as protein−ligand binding affinity and solubility.[63] The construction process involved importing inhibitor SMILES into the software and conducting 50 independent random splits with 105 inhibitors for the training set (80%) and 27 inhibitors for the test set (20%). Various Canvas fingerprints (radial, linear, dendritic, molprint2D) and descriptors (topographical, physicochemical, LigFilter) were extracted. Models were built using multiple linear regression (MLR), partial least-squares (PLS), kernel-based partial least-squares (KPLS), and principal component regression (PCR), with feature selection to identify the subset with all correlation coefficients below a specified threshold. Descriptors with more than 90% of structures having the same values were eliminated. The top 10-ranked QSKR models were retained. KPLS used both descriptor and fingerprint features, while other methods were trained using only descriptor features. In addition, we used the RDKit toolkit[64] to extract extended connectivity fingerprints 6 (ECFP6).

Feature Extraction Based on Molecular 3D Structures. Extracting features from molecular three-dimensional (3D) structures to build QSKR models provides a better understanding of the relationships between binding kinetics, stereochemical properties, stereoelectronic effects, hydrophobicity, and other properties. We used the Field-Based QSAR module in the Maestro software (Schrödinger, 2023) for feature extraction and model construction of N-HSP90α inhibitors' 3D structures. The process involved aligning the complex structures of 132 N-HSP90α inhibitors based on protein residues, separating the ligand's 3D structure, and using the Field-Based QSAR module for feature extraction and model construction. The data set was randomly split into a training set (80%) and a test set (20%), generating Gaussian steric, electrostatic, hydrophobic, hydrogen bond acceptor, and donor fields with a grid spacing of 1 Å. Grid points more than 2 Å away from any atom in the training set were excluded. Before PLS regression, each field attribute value was scaled, with a maximum of 4 PLS factors, and cross-validation was performed using the leave-one-out method. The data set underwent 10 different random splits, resulting in 10 different models. Additionally, 431 three-dimensional

https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645

descriptor features were extracted using the PaDEL package in the ChemDes platform.[65]

Feature Extraction Based on Complex Structures. We adopted the approach reported by Wang et al.[66] and employed the open-source generate_features.py script to extract features depicting interactions between ligand atoms and receptor residues in complex structures. This method characterizes protein−ligand interactions by defining N consecutive shells around each atom of the ligand and calculating the number of contacts between any ligand atom and any protein residue within those shells. We refer to this method as distance shell feature extraction (DSFE). Ligand atoms are categorized into eight types: C, H, O, N, P, S, HLA, and DU, where HLA represents halogen elements (F, Cl, Br, and I) and DU represents other elements. Meanwhile, the protein includes 20 standard amino acid residues and 1 nonstandard residue. Each shell has 168 possible combinations. Previously employed for extracting features of protein−ligand interactions, this method, combined with deep convolutional neural networks, achieved higher prediction accuracy for binding free energy and affinity compared to most traditional scoring functions. For the 132 complex structures, water molecules and other ions were removed. Ligands were uniformly named "LIG", saved in PDB format, and features were extracted using the generate_features.py script with 62 shells covering a range of 61.5 Å around each ligand atom, resulting in a feature matrix of 132 × 10,416.

Construction of the QSKR Models. AutoQSAR and Field-Based QSAR models were developed using Maestro software (Schrödinger, 2023), selecting models that performed best on the test set without overfitting. The machine learning training data set for the remaining methods consisted of extracted features and dissociation rate constant. The features underwent Min−Max normalization, scaling them to the range [0, 1]. eXtreme Gradient Boosting (XGBoost) and Support Vector Machine (SVM) regression were employed to assess each set of extracted features. These methods were chosen for their balanced and stable performance in comparison to other techniques such as Random Forest and Partial Least-Squares. The data set was randomly divided into a training set (80%) and a test set (20%). Feature selection was conducted on the training set based on feature importance thresholds. Hyperparameter tuning was performed using an exhaustive grid search with a 5-fold cross-validation, optimizing parameters that include the number of decision trees, learning rate, tree depth, and regularization parameter. During the formal training phase, the data set underwent 10 random splits, and leave-one-out cross-validation was applied to the training set (80%). Subsequently, the models were tested on the respective test sets.

Lead Compound Screening

To comprehensively evaluate the predictive accuracy of models constructed with different features and apply them to the development of novel N-HSP90α inhibitors, we employed a strategy that combines docking-based affinity screening with binding kinetics screening to identify potential small molecule inhibitors. Molecular docking is a computational technique to predict the binding mode and binding affinity of ligand and receptor proteins, which is widely used in drug discovery and design.[68−70] The screening process is succinctly outlined in Figure 2. Initially, structures of N-HSP90α residues 104−114 with four distinct conformations (PDB IDs: 5J2X, 1YC1, 2XHX, and 5LQ9) were optimized using the Protein Preparation module in the Maestro software (Schrödinger, 2023). Subsequently, the Receptor Grid Generation module was used to generate docking boxes for ligand binding. Finally, a docking-based screening was performed using the GVSrun (https://github.com/Wang-Lin-boop/Schrodinger-Script) script on a Linux cluster, using the Specs compound library (www.specs.net). The Specs compound library, comprising over 300,000 commercially available small molecules with proven performance, is widely used in lead compound screening. After docking, the top 4000 compounds based on Glide docking scores for each conformation were retained. Subsequently, the MM-GBSA module of the Maestro software (Schrödinger, 2023) was used to calculate the binding free energy for the top 1000 complexes in each conformation. The top 200 small molecules with the lowest binding free energy were then retained. Finally, 100 small molecules were selected based on scaffold similarity and binding free energy. Here, we used AutoQSAR models constructed within the Maestro software (Schrödinger, 2023) to predict the dissociation rate constant for these 100 small molecules, as this computational platform is widely employed in drug discovery and has demonstrated reliable predictive performance. Ultimately, considering the predicted dissociation rate constant and binding free energies, 39 small molecules were selected for experimental validation.

Protein Purification, Crystallization, and Structure Determination

The N-HSP90α protein expression and purification methods were consistent with those previously reported.[53] The cDNA sequence encoding human N-HSP90α WT (amino acid sequences 9−236) was inserted into the pET15b vector, incorporating an N-terminal His-TEV tag, and expressed in *Escherichia coli* BL21 (DE3). Cells were collected, resuspended in lysis buffer (2 × PBS, 20 mM imidazole, 10% glycerol, 1 mM TCEP, pH 7.5), and subjected to purification via a nickel column, followed by elution with 200 mM imidazole. Further purification was carried out using Superdex 200 (GE Healthcare) in 20 mM HEPES, 150 mM NaCl, 1 mM TCEP, pH 7.5.

Crystal structures with "loop-middle" and "loop-in" conformations were obtained in a crystallization buffer containing 100 mM bis-tris at pH 8.5, 20% PEG 3350, and 200 mM NaF (PDB IDs: 8W8K, 8KI4). However, the crystal structure in the "loop-in" conformation was obtained under conditions of 100 mM tris-HCl at pH 8.5, 200 mM $MgCl_2$, and 30% PEG 4000 (PDB ID: 8W4V). N-HSP90α WT protein was concentrated to 25 mg/mL, mixed in equal volume (1:1 v/v) with crystallization buffer, and grown at 4 °C for 3 days by hanging drop method. The crystals were immersed in a crystallization buffer containing a compound concentration of 5 mM. The sample was rapidly frozen in liquid nitrogen with 25% glycerol as a cryoprotectant. For structure determination and refinement, X-ray diffraction data were collected at the BL18U1 beam at the Shanghai Synchrotron Radiation Facility (SSRF). After HKL3000 treatment, the N-HSP90α *apo* structure was used as the search model to solve the problem using the phaser molecular replacement algorithm within the CCP4 software package. Subsequently, this model underwent further refinement using Coot and phenix. The data collection and refinement statistics for the solution structure are shown in Table S3.

Surface Plasmon Resonance (SPR) Analysis

The binding kinetic parameters of the selected compounds were measured by the surface plasmon resonance (SPR) method as described previously.[53] In brief, the recombinant N-HSP90α protein (10 mM HEPES pH 7.40, 0.15 M NaCl) was incubated with 100-fold alvespimycin (17-DMAG) hydrochloride and diluted to 20 μg/mL in sodium acetate pH 4.50. According to the Biacore standard, it was immobilized on the CM5 sensor chip (Cytiva) of the S series using amino coupling at a temperature of 25 °C and a flow rate of 10 μL/min, and the fixation amount was about 4000 RU. The unmodified carboxy-glucan matrix was used as a reference surface. The selected compounds were stored in 10 mM stock solution in 100% dimethyl sulfoxide (DMSO). The compounds were dissolved in the running buffer (20 mM HEPES pH 7.50, 150 mM NaCl, 0.05% Tween 20, 1 mM DTT, 0.1 mM EDTA, 2% DMSO) and then diluted in a 2-fold concentration gradient. Kinetic tests were performed at 25 °C with a flow rate of 30 μL/min, sample contact times of 120 s, and dissociation times of 200−600 s. Each cycle was cleaned with 50% DMSO, and solvent correction with different DMSO concentrations (1.5 to 2.5%) was performed every 48 cycles.

The data set was processed and analyzed within the Biacore Insight Evaluation Software (Cytiva). The solvent correction and two-reference association and dissociation data were fitted to a two-state reaction model fit with mass transport restrictions. This choice was motivated by the presence of conformational changes during the binding of compounds to N-HSP90α and the unsatisfactory outcomes obtained from fitting a simple 1:1 Langmuir binding model. To verify the two-state reaction model, different flow rates (20−30 μL/min) and prolonged binding times (120−200s) were tested to obtain saturation curves. The two-state reaction model describes the 1:1 binding of

https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645

Table 1. Validation Results of the Different Machine Learning Models of Dissociation Constant Rate ($\log(k_{off})$) Prediction[a]

structure	feature extraction	ML model	R^2	MAE
2D structure (SMILES)	AutoQSAR	KPLS	0.80	0.32
	ECFP6	XGBoost	0.777 ± 0.027	0.359 ± 0.026
		SVR	0.914 ± 0.094	0.174 ± 0.104
3D structure	Field-Based QSAR	PLS	0.74	0.41
	PaDEL_3D	XGBoost	0.875 ± 0.101	0.200 ± 0.098
		SVR	0.825 ± 0.075	0.305 ± 0.072
complex structure	Distance Shell Feature Extraction (DSFE)	XGBoost	0.812 ± 0.093	0.298 ± 0.072
		SVR	0.933 ± 0.065	0.182 ± 0.078

[a] R^2 is the coefficient of determination on the test set, and MAE is the mean absolute error between the observed and predicted values. The mean ± SD value was calculated by randomly splitting the data set 10 times. The AutoQSAR and Field-Based QSAR models were constructed using Maestro software (Schrödinger, 2023), and only the best-performing model was selected; thus, there is no standard deviation (SD).

injected analyte A and immobilized ligand B, followed by a conformational change of complex AB* to a more stable complex AB, as described in the following equations

$$[A] + [B] \underset{k_{d1}}{\overset{k_{a1}}{\Leftrightarrow}} [AB^*] \underset{k_{d2}}{\overset{k_{a2}}{\Leftrightarrow}} [AB]$$

$$K_D = 1/\left[\left(\frac{k_{a1}}{k_{d1}}\right)\times\left(1 + \frac{k_{a2}}{k_{d2}}\right)\right]$$

$$k_{off} = (k_{d1} \times k_{d2})/(k_{d1} + k_{a2} + k_{d2})$$

[71,72]

Circular Dichroic Spectrum

The CD spectra of N-HSP90α protein were obtained using a circular dichroic spectrometer (Chirascan Plus, UK) in a quartz cell with an optical path length of 1 mm. The N-HSP90α protein and excess compounds were incubated overnight at 4 °C, and the buffer was subsequently changed to 20 mM phosphate buffer at pH 7.5 to avoid interference from DMSO and chloride ions during the assay. In the final test, the protein concentration was adjusted to 0.15 mg/mL, the measurement temperature was 25 °C, the wavelength range was 180−260 nm, and the bandwidth was 1 nm. Each test consisted of three scans, and the results were averaged to minimize noise, with smoothing applied 3 degrees. To eliminate background baseline signals, blank experiments were conducted using the same buffer during testing.

All CD data were presented in Milli-Degrees at the time of testing, and the secondary structure was calculated using the instrument's built-in software CDNN. A neural network-based software as abbreviated as CDNN is a widely used program for analyzing circular dichroism (CD) data. It uses a reference database to deconvolute CD spectra and calculate the content of five different secondary structures.[73−75]

Isothermal Titration Calorimetry

All measurements were performed by PEAQ-ITC (Malvern) in buffer (20 mM HEPES, 150 mM NaCl, 1 mM TCEP, pH 7.5) with stirring at 750 rpm. The stock solution of the compound (40 mM) was diluted to 400 μM with ITC buffer, and the protein was diluted to 25 μM, both with a final DMSO concentration of 1%. During titration, 0.4 μL was injected, followed by 19 injections of the same 2 μL, each with a duration of 4 and 120 s between each injection. The acquired data were processed using the PEAQ-ITC analysis software.

MD Simulations

A coarse-grained Cα model was used to simulate the conformational changes of *apo* N-HSP90α. Initially, structures for the "loop-middle" (PDB ID: 8KI4), "loop-in" (PDB ID: 1YER), and "loop-out" (PDB ID: 1YES) conformations within residues 104−114 were processed using the Protein Preparation module in Maestro software (Schrödinger, 2023). This process involved determining the protonation states of protein residues under pH 7.4 conditions, optimizing hydrogen bond networks, and performing energy minimization. Subsequently, Cα structures were further generated using the SMOG2 (version 2.5-beta) template (SBM_CA). SMOG2 is packaged with templates for some commonly used structure-based Hamiltonians, defining a specific force field (more details on the SBM_CA template can be found in the Supporting Information (SI)).[76−78] GROMACS (v5.0.4) and PLUMED (v2.3.8) were used to conduct metadynamics simulations on different conformations of *apo* N-HSP90α.[79] GROMACS and PLUMED are widely used software for molecular dynamics simulations, with PLUMED capable of employing different enhanced sampling algorithms and collective variables. (CVs).[80,81] The Langevin equation was used to run simulations in a low-friction limit. Simulation of the three conformations for an equivalent duration of 1 μs maintained identical simulation conditions and parameters (temperature 310 K). The root-mean-square deviation (RMSD) of Cα 102−137 served as the collective variable, and a Gaussian bias potential energy was introduced every 0.5 ps, with a Gaussian width of 0.05 nm and a height of 1 kJ/mol. When the root-mean-square deviation (RMSD) exceeds 0.5 nm, a bias potential energy is introduced to restore system equilibrium. The three trajectories were combined for postsimulation analysis, and the free-energy landscape was extracted using GROMACS' integrated tools. The "loop-out" conformation served as the reference structure, with the RMSD of the collective variable and Cα 109 distance employed as reaction coordinates.

RESULTS AND DISCUSSION

The quality metrics of the regression model for the dissociation rate constant ($\log(k_{off})$) of N-HSP90α are presented in Table 1. Both the AutoQSAR model constructed in Maestro software and the Field-based QSAR model exhibit excellent predictive capabilities on the test set. Within the Field-based QSAR model, the primary contributors are identified as the Gaussian steric field and the Gaussian hydrophobic field. This highlights the significance of inhibitors occupying the hydrophobic pocket of the ATP-binding site and engaging in extensive hydrophobic interactions in influencing the k_{off}. Moreover, models constructed with features extracted from molecular 2D structures, 3D structures, and complex structures also exhibit superior predictive performance. Features extracted from molecular 3D structures indicate that variations in molecular polarity, nonpolarity, van der Waals volumes, and other factors significantly contribute to these models. It is noteworthy that the model constructed using the DSFE method to extract features from the complex has an R^2 of 0.93 and an MAE of 0.18 on the test set. The top 10 features contributing the most to the model primarily involve interactions between tryptophan, leucine, and the nitrogen atoms on the ligand, as well as interactions between asparagine and hydrogen atoms. This highlights the significant roles played by hydrophobic interactions and hydrogen bonds in the process of ligand dissociation.

However, when dividing the data set based on structural scaffolds, with hydroxy-indazole scaffold inhibitors and

https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645

Table 2. Experimentally Determined Binding Kinetics for 16 Compounds, with Predicted Values for k_{off} from the AutoQSAR Model[a]

	N-HSP90α-WT				
compound IDs	K_D (M)	k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	log (k_{off})	log (k_{off})-Pre
1	6.05×10^{-6}	1.38×10^{3}	5.31×10^{-3}	−2.2747	−1.66
2	5.47×10^{-6}	4.21×10^{3}	2.31×10^{-2}	−1.6359	−1.863
3	2.32×10^{-5}	2.86×10^{3}	3.79×10^{-2}	−1.4208	−2.013
4	5.87×10^{-6}	3.13×10^{3}	1.51×10^{-2}	−1.8206	−1.563
5	6.06×10^{-6}	4.71×10^{4}	1.17×10^{-1}	−0.9321	−1.712
6	5.42×10^{-6}	4.69×10^{3}	1.23×10^{-2}	−1.7588	−1.869
7	4.62×10^{-5}	2.75×10^{2}	1.02×10^{-2}	−1.9901	−2.535
8	1.23×10^{-5}	1.48×10^{3}	8.74×10^{-3}	−2.0586	−1.828
9	1.33×10^{-5}	3.45×10^{3}	4.84×10^{-2}	−1.3149	−2.041
10	2.98×10^{-5}	1.31×10^{2}	4.81×10^{-3}	−2.3177	−1.746
11	2.20×10^{-5}	4.53×10^{1}	9.23×10^{-4}	−3.0346	−1.93
12	3.75×10^{-5}	3.89×10^{2}	1.39×10^{-2}	−1.8565	−1.704
13	6.87×10^{-5}	3.05×10^{1}	2.14×10^{-3}	−2.6702	−2.351
14	1.77×10^{-5}	1.09×10^{3}	1.91×10^{-2}	−1.7192	−2.266
15	1.64×10^{-5}	1.71×10^{3}	1.93×10^{-2}	−1.7144	−2.286
16	1.78×10^{-5}	1.19×10^{3}	2.30×10^{-2}	−1.6375	−2.546

[a]Data were averaged from 2 to 4 independent measurements, and the corresponding standard deviations are presented in Table S2. "log(k_{off})-Pre" indicates the predicted values of the AutoQSKR model.

resorcinol scaffold inhibitors as the training set (102 compounds, ∼77.2%), and other scaffold inhibitors as the test set (30 compounds, ∼22.8%), these models display limited generalization ability on the test set (the best-performing DSFE_SVR model has an R^2 of 0.34 and an MAE of 0.979). This limitation may be attributed to an insufficient data set or an imbalance in data distribution, resulting in a failure to adequately capture the differences in inhibitor features. When inhibitors of helical conformation were used as the training set (108 compounds, ∼81.8%), the prediction accuracy of the DSFE_SVR model on the test set (inhibitors of loop conformation) decreased (R^2 = 0.40, MAE = 0.545). This highlights the differences in determining factors for dissociation rate constants between inhibitors with helical and loop conformations. In conclusion, these findings underscore the critical role of diverse structural and interaction features in predicting the dissociation rate constant of N-HSP90α inhibitors.

Experimental Validation of N-HSP90α Compound Screening

Here, we performed surface plasmon resonance (SPR) experiments to test the selected 39 compounds. Among these, 31 compounds exhibited binding signals in steady-state affinity assays, while 16 compounds showed stable association and dissociation responses in binding kinetics assays. This allowed us to obtain the respective binding kinetics parameters, detailed in Table 2. The binding affinity of compounds with different scaffolds ranged from a few micromoles (μM) to several tens of micromoles (μM), and some compounds displayed relatively slow dissociation rate constant. The chemical structures of these 16 inhibitors are illustrated in Figure S4, and they exhibit relatively low similarity to the 132 compounds in the data set (Figure S5). Notably, we used a two-state reaction model to fit the binding kinetic parameters. This model describes a 1:1 binding of analyte to immobilized ligand followed by a conformational or other changes that stabilize the complex. This choice is motivated by the high flexibility and plasticity of the *apo* N-HSP90α's ATP-lid, which undergoes conformational changes upon binding different compounds, making the binding and dissociation processes more complex. Binding and dissociation curves suggest that the compound initially forms an intermediate state AB* with the N-HSP90α protein, followed by conformational rearrangement to form the more stable complex AB (Figure S6). Additionally, in this study, we further employed circular dichroism experiments to validate ligand-induced N-HSP90α conformational changes (Figure S7). Upon the addition of compounds **14**, a significant increase of 2.8% in α-helix content was observed relative to the *apo* structure, representing approximately 8.2% of the total α-helix content. This indicates that these compounds bind to the target and induce the rearrangement of residues 104−114 into a helical structure. Guldenhaupt et al. also verified secondary structure changes of the HSP90 protein induced by resorcinol inhibitors with different substituents by attenuated total reflection infrared spectroscopy (ATR-FTIR).[82] Table 2 presents the predicted log(k_{off}) values using the AutoQSAR model and the experimentally measured log(k_{off}) values for these 16 compounds. In comparison, the log(k_{off}) prediction deviations for most compounds are within 1 order of magnitude.

Through surface plasmon resonance (SPR) experiments validating the dissociation rate constant, we further assessed the practical predictive performance of each regression model. The results indicate that, for these 16 new scaffolds of N-HSP90α inhibitors, most models exhibit mean absolute errors (MAEs) lower than an order of magnitude (Figure 3). Among them, the AutoQSAR_KPLS model stands out, with a remarkable MAE of 0.516 for predicting the dissociation rate constant (log(k_{off})). This model was created using linear fingerprints in Maestro software, selected as the optimal model through 50 random data set splits. Figure S8 shows the regression results of the ECFP6_SVR model. Compared to the QSKR models established using Maestro software, the ECFP6_SVR and DSFE_SVR models demonstrate better predictive performance on the test set (R^2 > 0.9) and similarly impressive predictive performance on the 16 new scaffold compounds. However, models constructed by extracting features from the 3D structure of molecules exhibit lower performance on the experimental

https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645

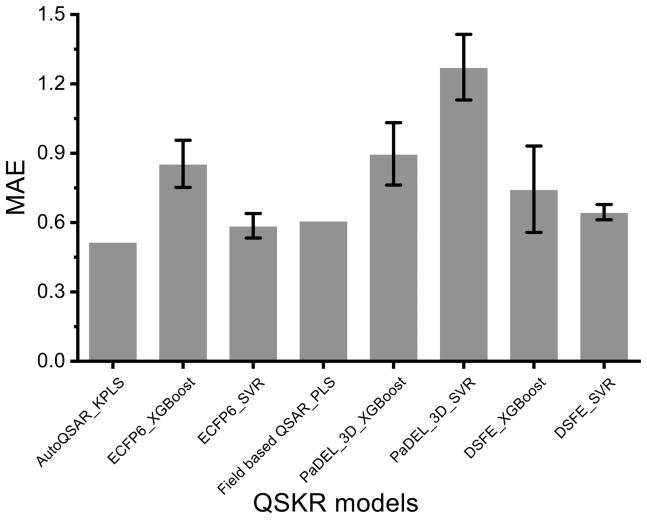


Figure 3. Mean absolute error (MAE) between predicted and experimental values. The error bars represent the standard deviation (SD) of MAE, which was estimated based on 10 random splits of the data set.

validation set, potentially due to a significant contribution of molecular weight in the extracted features. Additionally, we performed the Wilcoxon signed-rank test to compare the predicted values of these models with experimental observations (Table S2). The results indicate that, at a significance level of 0.05, there is no significant difference between the predicted and observed values. This further validates the reliability and accuracy of the models. In conclusion, these models show promising predictive potential, providing robust tools for the discovery of new HSP90 inhibitors.

Crystal Structures of Bound Forms and Ligand-Free of N-HSP90α

To delve into the relationship between different conformations and the binding kinetics of inhibitors, we conducted crystallization experiments and determined the complex structure of compound **4** with N-HSP90α in the "loop-in" conformation (Figure 4A). Consistent with the binding mode of previously reported resorcinol derivatives, the 2-hydroxy group on the resorcinol ring of compound **4** forms a hydrogen bond with the D93 side chain, and it engages in water-mediated hydrogen bonding within the hydrophobic pocket with G92. Additionally, the 4-hydroxy group interacts via hydrogen bonding with N51 and forms water-mediated hydrogen bonds

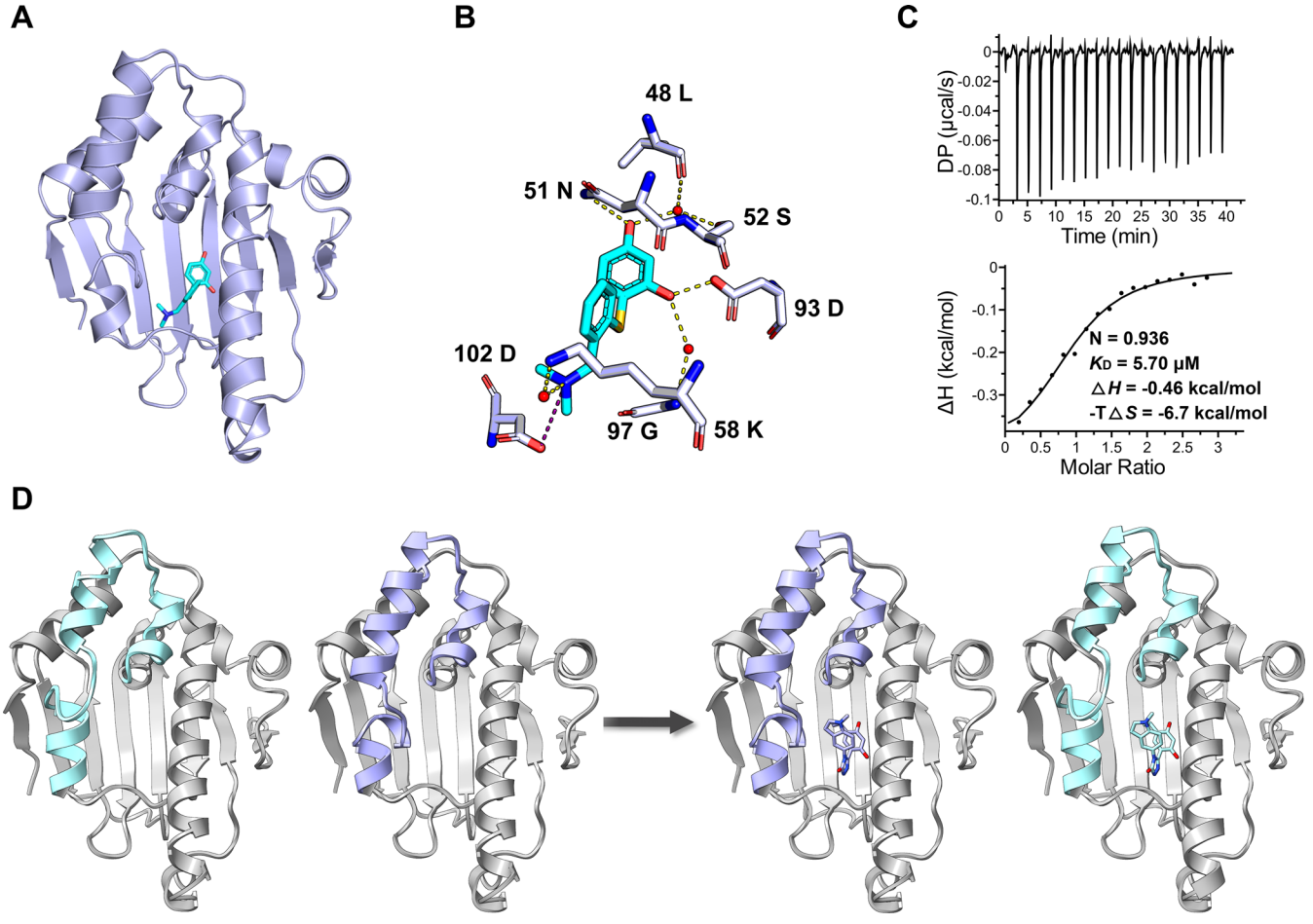


Figure 4. Different conformational complex structures of N-HSP90α. (A) Cartoon representation of the complex structure of compound **4** with N-HSP90α in the "loop-in" conformation (PDB ID: 8W4V). (B) Detailed interaction of compound **4** with relevant residues of N-HSP90α and water molecules in the ligand binding pocket. Yellow represents hydrogen bonds and purple represents salt bridges. (C) Thermodynamics data of compound **4** titrated into N-HSP90α with a concentration ratio of 400 μM: 25 μM measured by PEAQ-ITC. (D) In one asymmetric unit, both the *apo* form and the Ganetespid-bound form of N-HSP90α exhibit "loop-in" and "loop-middle" conformations (PDB IDs: 8KI4, 8W8K).

https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645

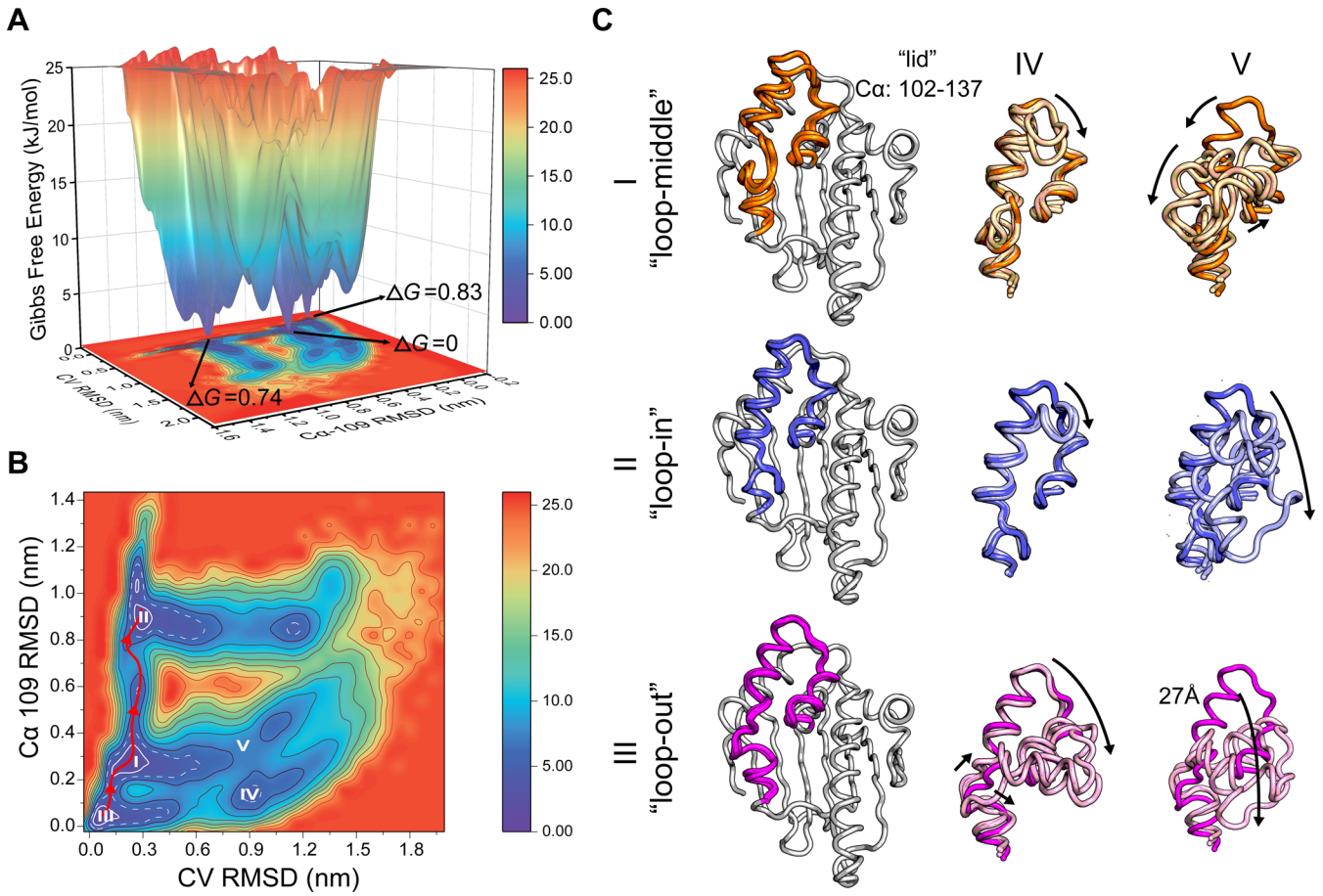


Figure 5. Free-energy landscape and structural analysis of *apo* N-HSP90α metadynamics simulations in different conformations. (A) 3D plot illustrating the free-energy landscape obtained from the merged trajectories. The figure displays the free energies associated with the three lowest-energy basins. (B) A 2D free-energy topography was constructed using the root-mean-square deviation (RMSD) of the collective variable, Cα 102−137, and the RMSD of Cα 109 as reaction coordinates. Regions I, II, and III correspond to a free-energy range of 0−3.05 kJ/mol (solid white line area), while region IV represents a free-energy range of 3.05−5.42 kJ/mol (dotted white line area). Region V indicates a free energy greater than 5.42 kJ/mol. The red solid line represents the lowest-energy path for the three conformations I, II, and III allosteric to each other. (C) Structural analysis of regions I, II, III, IV, and V was conducted following clustering. The structure only displays the Cα 102−137 segment's "lid", while other regional conformations remain unchanged. The black arrow denotes the direction of "lid" movement during the simulation.

with L48 and S52. The hydrogen atom on amino group 15 participates in water-mediated hydrogen bonding with K58, while amino group 15 establishes a salt bridge with the D102 residue (Figure 4B). We also determined the thermodynamic parameters for compounds **4** and **5** through ITC (Figures 4C and S9).

In X-ray crystallography experiments, the "loop-middle" conformation of *apo* N-HSP90α was observed for the first time (previously, the "loop-middle" conformation had only been observed in *holo*-N-HSP90α structures). Interestingly, we observed a structure in one asymmetric unit where *apo* N-HSP90α simultaneously adopts both "loop-in" and "loop-middle" conformations (PDB ID: 8KI4, Figure 4D, left). Similarly, after binding with Ganetespid (an N-HSP90α inhibitor in phase III), these two conformations are also present (PDB ID: 8W8K, Figure 4D, right). This suggests that residues 104−114 of the ATP-lid in *apo* N-HSP90α exist in different loop conformations in solution. The binding of specific ligands, different crystallization conditions, additives, pH, and other factors can alter the pre-existing conformational equilibrium.[51,52] This leads to a shift in the population distribution toward more stable conformations, thereby increasing the proportion of this conformation. In the isolated structure of N-HSP90α, Henot et al. also observed a "closed" conformation of the ATP-lid, with an occupancy of 3−4% at room temperature.[52] However, this helical conformation remains elusive in the absence of ligand binding, likely triggered by specific binding events. Research by Amaral et al. on the thermodynamics of the N-HSP90α inhibitor binding indicates that inhibitors capable of forming helical conformations exhibit significantly favorable entropy contributions upon binding. Overcoming higher-energy barriers during binding and dissociation, these specific compounds dynamically alter the conformational landscape.[53] Compounds adept at occupying hydrophobic pockets and forming extensive hydrophobic interactions further induce the rearrangement of ATP-lid from different loop conformations into a helical conformation. Inhibitors with the ability to induce the formation of helical conformations exhibit slower rates of binding or dissociation and can linger in this conformation for an extended duration after binding to the protein. This may reduce the fraction of conformations suitable for normal chaperone functional cycling.

https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645

Molecular Dynamics Simulations of *apo* N-HSP90α in Different Conformations

The binding kinetics of N-HSP90α inhibitors are intricately linked to their distinct conformational states, and a comprehensive understanding of the entire conformational system is pivotal for elucidating the molecular mechanism of N-HSP90α. In recent research, molecular dynamics simulations have been extensively used to explore the mechanisms of HSP90 conformational changes and inhibitor binding.[83,84] Chen et al. demonstrated, through their study on the binding mechanisms of inhibitors capable of forming helical conformations, that inhibitor binding distinctly influences the structural flexibility of N-HSP90α, potentially leading to significant impacts on HSP90 function.[85,86] Conformational changes in the ATP-lid occur on the microsecond to millisecond time scale and even longer.[53] No significant changes in the ATP-lid were observed in our all-atom dynamics simulations of a few microseconds. To address this challenge, we used a coarse-grained model for metadynamics simulations. Onuchic and Wolynes have made significant contributions to the study of protein folding, energy landscapes, and biomolecular dynamics.[87−89] Structure-based Hamiltonians, simplified models are derived from the energy landscape theory of protein folding, which has proven effective in capturing long-time and large-length scale motions of proteins.[90−92] These techniques enable us to describe different conformational energy landscapes and observe structural changes, further enhancing our understanding of the structure−function relationship.

Using the coarse-grained Cα structures of "loop-in", "loop-out", and "loop-middle" in *apo* N-HSP90α as initial structures, the RMSD of Cα 102−137 served as a collective variable (Figure S10). A bias potential energy was applied to configurations with significant deviations from these conformations, facilitating the overcoming of high-energy barriers for transition states. Throughout the entire simulation process, all three initial conformations could effectively and relatively rapidly transition into other conformations. Through the analysis of combined trajectories, the RMSD of the collective variable and the distance between Cα atoms at position 109 as reaction coordinates are used to calculate the free-energy landscape. In the free-energy landscape (Figure 5A), the three basins with the lowest energy, denoted as ①, ②, and ③, correspond to the "loop-middle", "loop-in", and "loop-out" conformations, with the relative free energy of 0, 0.74, and 0.83 kJ/mol, respectively. Additionally, multiple potential wells around them represent the diversity of different loop conformations. Based on the analysis of their free-energy ranges and conformational clustering, we divided them into five regions (Figure 5B), with regions I, II, and III representing "loop-middle", "loop-in", and "loop-out" conformations, respectively. Within the range of 0−3.05 kJ/mol in free energy (Figure 5C), the ATP-lid is relatively stable. The IV region comprises metastable conformations with a free-energy range of 3.05−5.42 kJ/mol. In this "loop-out" state, the ATP-lid obviously closes toward the ATP pocket. Region V represents high-energy conformations, with free energies exceeding 5.42 kJ/mol. Notably, the ATP-lid segment of the three structures exhibits different degrees of conformational regulation in the high-energy region, with multiple "closed" conformations observed. This segment demonstrates significant flexibility while inducing changes in Cα 104−114 near the ATP pocket. Simulation results suggest the presence of a diverse distribution of cyclic conformations in *apo* N-HSP90α, with the major proportions being the "loop-middle", "loop-in", and "loop-out" conformations. In contrast, the ATP-lid of the "loop-out" conformation tends to close more easily, transitioning to other states. The red line in the free-energy landscape represents the lowest-energy conformational change path connecting "loop-out" and "loop-in" conformations (Figure 5B). The pathway search method used in the study is similar to previously reported methods,[35] coarsely dividing the constructed free-energy landscape into a 32 × 32 grid, with each grid having corresponding energy and coordinates. Starting from the lowest-energy point in the "loop-out" conformation, we traverse the grid along the path with the minimum global potential energy. This path includes three stable states and one metastable conformation, accompanied by three transitional conformations (Figure S11). With the opening and closing of the ATP-lid, residues 104−114 display different loop conformations. The differences in the size of the ATP-binding pocket and solvent distribution among different cyclic conformations may be due to significant changes in the ATP-lid of *apo* N-HSP90α. Ligand binding tends to prioritize the binding of the matching conformation, thereby increasing its proportion. In the simulations, the formation of the helical conformation was not observed, suggesting that the induction of helical conformation formation, triggered by specific ligand binding events, requires overcoming higher-energy barriers.

CONCLUSIONS

The role of drug−target binding kinetics in drug efficacy and safety is becoming increasingly important. The rapid and accurate prediction of the binding kinetics of drug−target has become a current research focus with the aim of simplifying and expediting the drug discovery process. In this study, we developed a quantitative structure−kinetics relationship (QSKR) model for the dissociation rate constant (k_{off}) of HSP90 inhibitors on a larger data set. The model exhibited good predictive performance on the test set and was successfully applied to a lead compound screening strategy, leading to the discovery of new N-HSP90α inhibitors. Consistent with previous research findings, the dissociation rate of inhibitors is closely related to hydrogen-bonding and hydrophobic interactions and factors such as the occupation of the hydrophobic pocket within the ATP-binding site. The model shows further refinement and extension potential on an experimental validation set with a new scaffold. Our experimental data and molecular dynamics simulations indicate that residues 104−114 of *apo* N-HSP90α ATP-lid exhibit different loop conformations. The transitions between different loop conformations are induced by the conformational changes of the ATP-lid. Inhibitor binding may involve a mixed mechanism of conformational selection and induced fit. The conformational dynamics of N-HSP90α are closely related to the binding and dissociation behavior of inhibitors. However, currently, most N-HSP90α inhibitors primarily compete by binding to ATP rather than being specifically designed to influence the conformation of N-HSP90α. Inhibitors with longer residence times and the ability to induce helical conformations may profoundly impact the conformational distribution of N-HSP90α, suggesting that small molecules influencing N-HSP90α conformational dynamics may have clinical potential.

In AI, different scientific problems or models require different amounts of data, which is directly related to the complexity of the problem itself. Let us take our example: since N-HSP90α has multiple dynamic conformations, which indicates that the process of small molecule recognition is also very complex, the

https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645

variable dimensions are high and therefore a larger number of parameters that can cover these different dimensions are needed to train the model. In our example, we used more than 130 data to get better results, which indicates that this data size is reasonably meaningful, but the generalization ability of the model still needs to be improved, for which more data is needed. Of course, in artificial intelligence problems, we can simplify the problem by creating a physical model and artificially reducing the dimensionality when the data is relatively small. There are currently some studies on the development of computational models for the prediction of binding kinetics. For example, Ganotra et al. used the comparative binding energy (COMBINE) method, calculating interaction energies for ligand−receptor complexes in data sets comprising 70 N-HSP90α inhibitors and 36 HIV-1 inhibitors. They constructed QSKR models using partial least-squares, achieving correlation coefficients of 0.69 and 0.70, and the mean absolute errors were 0.45 and 0.58 on two test data sets, respectively. This indicates the feasibility of constructing quantitative structure−kinetics relationship prediction models for other systems, even with a data set size below 100. In our study, a predictive model was developed using a data set of 132 N-HSP90α inhibitors, with a determination coefficient (R^2) of 0.93 and a mean absolute error (MAE) of 0.18. It is noteworthy that, when the data set is split according to scaffold types, a significant decline in performance on the test set was observed. This suggests that due to the limited data and imbalance in inhibitor types, the predictive accuracy for other scaffolds is reduced, and the generalization capability of the model needs improvement. Therefore, obtaining more binding kinetics data would be highly desirable. With the determination of more structures and binding kinetics data, AI-based models for binding kinetics prediction are expected to have greater advantages and prospects, providing substantial support for drug discovery and development.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.4c00123.

The SMILES and binding kinetics information on the 132 HSP90 inhibitors used in this study, as well as the SMILES of 39 compounds tested in SPR experiments mentioned in this article (PDF)

Accurate characterization of binding kinetics and allosteric mechanisms for the HSP90 chaperone inhibitors using AI-augmented integrative biophysical studies (Table S1) (XLSX)

Accurate characterization of binding kinetics and allosteric mechanisms for the HSP90 chaperone inhibitors using AI-augmented integrative biophysical studies (Table S2) (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Fang Bai − *Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology and School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China; Shanghai Clinical Research and Trial Center, Shanghai 201210, China;* orcid.org/0000-0003-1468-5568; Email: baifang@shanghaitech.edu.cn

Authors

Chao Xu − *Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China*

Xianglei Zhang − *Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China*

Lianghao Zhao − *Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China*

Gennady M. Verkhivker − *Keck Center for Science and Engineering, Graduate Program in Computational and Data Sciences, Schmid College of Science and Technology, Chapman University, Orange, California 92866, United States; Department of Biomedical and Pharmaceutical Sciences, Chapman University School of Pharmacy, Irvine, California 92618, United States;* orcid.org/0000-0002-4507-4471

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.4c00123

Author Contributions

#C.X. and X.Z. contributed equally to this work. F.B. designed the project. C.X. constructed and trained the QSKR models, expressed the protein, and performed the SPR experiment, crystal experiment, and molecular dynamics simulations. X.Z. collected crystal diffraction data and worked with C.X. to perform the analysis on the experimental studies; C.X., X.Z., L.Z., G.M.V., and F.B. analyzed the computational results. C.X. and F.B. wrote the manuscript. All authors read and approved the final manuscript. CRediT: **Chao Xu** formal analysis, investigation, validation, visualization, writing-original draft, writing-review & editing; **Xianglei Zhang** data curation, formal analysis, writing-review & editing; **Lianghao Zhao** data curation, methodology; **Gennady M. Verkhivker** visualization, writing-review & editing; **Fang Bai** conceptualization, funding acquisition, supervision, writing-original draft, writing-review & editing.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Key R&D Program of China (Grant IDs: 2022YFC3400501, 2022YFC3400504, and 2022YFC3400500), the Shanghai Science and Technology Development Funds (Grant ID: 22ZR1441400), the National Natural Science Foundation of China (Grant IDs: 82341093, 82003654), start-up package from ShanghaiTech University, and Shanghai Frontiers Science Center for Biomacromolecules and Precision Medicine at ShanghaiTech University. G.M.V. was supported by the National Institutes of Health under Award 1R01AI181600-01 and NIH Subaward 6069-SC24-11. The authors also thank the HPC Platform of ShanghaiTech University and the Discovery Technology Platform of SIAIS, ShanghaiTech University, for assistance and the staffs from BL18U1 beamline of the National Facility for Protein Science in Shanghai (NFPS) at Shanghai Synchrotron Radiation Facility for assistance during data collection.

REFERENCES

(1) Waring, M. J.; Arrowsmith, J.; Leach, A. R.; Leeson, P. D.; Mandrell, S.; Owen, R. M.; Pairaudeau, G.; Pennie, W. D.; Pickett, S. D.; Wang, J.; et al. An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat. Rev. Drug Discovery* **2015**, *14* (7), 475−486.

(2) Copeland, R. A. The drug-target residence time model: a 10-year retrospective. *Nat. Rev. Drug Discovery* **2016**, *15* (2), 87−95.

(3) Zhang, R.; Monsma, F. Binding kinetics and mechanism of action: toward the discovery and development of better and best in class drugs. *Expert Opin. Drug Discovery* **2010**, *5* (11), 1023−1029.

(4) Lu, H.; Tonge, P. J. Drug-target residence time: critical information for lead optimization. *Curr. Opin. Chem. Biol.* **2010**, *14* (4), 467−474.

(5) Vauquelin, G. Effects of target binding kinetics on in vivo drug efficacy: koff, kon and rebinding. *Br. J. Pharmacol.* **2016**, *173* (15), 2319−2334.

(6) Bernetti, M.; Masetti, M.; Rocchia, W.; Cavalli, A. Kinetics of Drug Binding and Residence Time. *Annu. Rev. Phys. Chem.* **2019**, *70*, 143−171.

(7) Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug-target residence time and its implications for lead optimization. *Nat. Rev. Drug Discovery* **2006**, *5* (9), 730−739.

(8) Schuetz, D. A.; de Witte, W. E.; Wong, Y. C.; Knasmueller, B.; Richter, L.; Kokh, D. B.; Sadiq, S. K.; Bosma, R.; Nederpelt, I.; Heitman, L. H.; Segala, E.; Amaral, M.; Guo, D.; Andres, D.; Georgi, V.; Stoddart, L. A.; Hill, S.; Cooke, R. M.; De Graaf, C.; Leurs, R.; Frech, M.; Wade, R. C.; de Lange, E. C. M.; IJzerman, A. P.; Müller-Fahrnow, A.; Ecker, G. F. Kinetics for Drug Discovery − An industry driven effort to target drug residence time. *Drug Discovery Today* **2017**, *22* (6), 896−911.

(9) Tummino, P. J.; Copeland, R. A. Residence time of receptor-ligand complexes and its effect on biological function. *Biochemistry* **2008**, *47* (20), 5481−5492.

(10) Guo, D.; Mulder-Krieger, T.; AP, I. J.; Heitman, L. H. Functional efficacy of adenosine A(2)A receptor agonists is positively correlated to their receptor residence time. *Br. J. Pharmacol.* **2012**, *166* (6), 1846−1859.

(11) Barnes, P. J. Tiotropium bromide. *Expert Opin. Invest. Drugs* **2001**, *10* (4), 733−740.

(12) Guo, D.; Hillger, J. M.; AP, I. J.; Heitman, L. H. Drug-target residence time–a case for G protein-coupled receptors. *Med. Res. Rev.* **2014**, *34* (4), 856−892.

(13) Pan, A. C.; Borhani, D. W.; Dror, R. O.; Shaw, D. E. Molecular determinants of drug-receptor binding kinetics. *Drug Discovery Today* **2013**, *18* (13−14), 667−673.

(14) Tautermann, C. S. Impact, determination and prediction of drug-receptor residence times for GPCRs. *Curr. Opin. Pharmacol.* **2016**, *30*, 22−26.

(15) Schuetz, D. A.; Richter, L.; Amaral, M.; Grandits, M.; Gradler, U.; Musil, D.; Buchstaller, H. P.; Eggenweiler, H. M.; Frech, M.; Ecker, G. F. Ligand Desolvation Steers On-Rate and Impacts Drug Residence Time of Heat Shock Protein 90 (Hsp90) Inhibitors. *J. Med. Chem.* **2018**, *61* (10), 4397−4411.

(16) Casasnovas, R.; Limongelli, V.; Tiwary, P.; Carloni, P.; Parrinello, M. Unbinding Kinetics of a p38 MAP Kinase Type II Inhibitor from Metadynamics Simulations. *J. Am. Chem. Soc.* **2017**, *139* (13), 4780−4788.

(17) Wang, J.; Do, H. N.; Koirala, K.; Miao, Y. Predicting Biomolecular Binding Kinetics: A Review. *J. Chem. Theory Comput.* **2023**, *19* (8), 2135−2148.

(18) Tiwary, P.; Limongelli, V.; Salvalaglio, M.; Parrinello, M. Kinetics of protein-ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112* (5), E386−391.

(19) Tiwary, P.; Mondal, J.; Morrone, J. A.; Berne, B. J. Role of water and steric constraints in the kinetics of cavity-ligand unbinding. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (39), 12015−12019.

(20) Tiwary, P.; Mondal, J.; Berne, B. J. How and when does an anticancer drug leave its binding site? *Sci. Adv.* **2017**, *3* (5), No. e1700014.

(21) Donyapour, N.; Roussey, N. M.; Dickson, A. REVO: Resampling of ensembles by variation optimization. *J. Chem. Phys.* **2019**, *150* (24), No. 244112.

(22) Lotz, S. D.; Dickson, A. Wepy: A Flexible Software Framework for Simulating Rare Events with Weighted Ensemble Resampling. *ACS Omega* **2020**, *5* (49), 31608−31623.

(23) Nunes-Alves, A.; Zuckerman, D. M.; Arantes, G. M. Escape of a Small Molecule from Inside T4 Lysozyme by Multiple Pathways. *Biophys. J.* **2018**, *114* (5), 1058−1066.

(24) Dixon, T.; Lotz, S. D.; Dickson, A. Predicting ligand binding affinity using on- and off-rates for the SAMPL6 SAMPLing challenge. *J. Comput. Aided Mol. Des.* **2018**, *32* (10), 1001−1012.

(25) Plattner, N.; Noe, F. Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. *Nat. Commun.* **2015**, *6*, No. 7653.

(26) Mondal, J.; Ahalawat, N.; Pandit, S.; Kay, L. E.; Vallurupalli, P. Atomic resolution mechanism of ligand binding to a solvent inaccessible cavity in T4 lysozyme. *PLoS Comput. Biol.* **2018**, *14* (5), No. e1006180.

(27) Mollica, L.; Theret, I.; Antoine, M.; Perron-Sierra, F.; Charton, Y.; Fourquez, J. M.; Wierzbicki, M.; Boutin, J. A.; Ferry, G.; Decherchi, S.; et al. Molecular Dynamics Simulations and Kinetic Measurements to Estimate and Predict Protein-Ligand Residence Times. *J. Med. Chem.* **2016**, *59* (15), 7167−7176.

(28) Sinko, W.; Miao, Y.; de Oliveira, C. A.; McCammon, J. A. Population based reweighting of scaled molecular dynamics. *J. Phys. Chem. B* **2013**, *117* (42), 12759−12768.

(29) Bianciotto, M.; Gkeka, P.; Kokh, D. B.; Wade, R. C.; Minoux, H. Contact Map Fingerprints of Protein-Ligand Unbinding Trajectories Reveal Mechanisms Determining Residence Times Computed from Scaled Molecular Dynamics. *J. Chem. Theory Comput.* **2021**, *17* (10), 6522−6535.

(30) Nunes-Alves, A.; Kokh, D. B.; Wade, R. C. Ligand unbinding mechanisms and kinetics for T4 lysozyme mutants from tauRAMD simulations. *Curr. Res. Struct. Biol.* **2021**, *3*, 106−111.

(31) Kokh, D. B.; Amaral, M.; Bomke, J.; Gradler, U.; Musil, D.; Buchstaller, H. P.; Dreyer, M. K.; Frech, M.; Lowinski, M.; Vallee, F.; et al. Estimation of Drug-Target Residence Times by tau-Random Acceleration Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2018**, *14* (7), 3859−3869.

(32) Yang, Y. I.; Shao, Q.; Zhang, J.; Yang, L.; Gao, Y. Q. Enhanced sampling in molecular dynamics. *J. Chem. Phys.* **2019**, *151* (7), No. 070902.

(33) Lazim, R.; Suh, D.; Choi, S. Advances in Molecular Dynamics Simulations and Enhanced Sampling Methods for the Study of Protein Systems. *Int. J. Mol. Sci.* **2020**, *21* (17), No. 6339.

(34) Bai, F.; Xu, Y.; Chen, J.; Liu, Q.; Gu, J.; Wang, X.; Ma, J.; Li, H.; Onuchic, J. N.; Jiang, H. Free energy landscape for the binding process of Huperzine A to acetylcholinesterase. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (11), 4273−4278.

(35) Bai, F.; Jiang, H. Computationally Elucidating the Binding Kinetics for Different AChE Inhibitors to Access the Rationale for Improving the Drug Efficacy. *J. Phys. Chem. B* **2022**, *126* (40), 7797−7805.

(36) Qu, S.; Huang, S.; Pan, X.; Yang, L.; Mei, H. Constructing Interconsistent, Reasonable, and Predictive Models for Both the Kinetic and Thermodynamic Properties of HIV-1 Protease Inhibitors. *J. Chem. Inf. Model* **2016**, *56* (10), 2061−2068.

(37) Schuetz, D. A.; Richter, L.; Martini, R.; Ecker, G. F. A structure-kinetic relationship study using matched molecular pair analysis. *RSC Med. Chem.* **2020**, *11* (11), 1285−1294.

(38) Ganotra, G. K.; Wade, R. C. Prediction of Drug-Target Binding Kinetics by Comparative Binding Energy Analysis. *ACS Med. Chem. Lett.* **2018**, *9* (11), 1134−1139.

(39) Kokh, D. B.; Kaufmann, T.; Kister, B.; Wade, R. C. Machine Learning Analysis of τRAMD Trajectories to Decipher Molecular Determinants of Drug-Target Residence Times. *Front. Mol. Biosci.* **2019**, *6*, No. 36.

https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645

(40) Sohraby, F.; Nunes-Alves, A. Advances in computational methods for ligand binding kinetics. *Trends Biochem. Sci.* **2023**, *48* (5), 437−449.
(41) Amangeldiuly, N.; Karlov, D.; Fedorov, M. V. Baseline Model for Predicting Protein−Ligand Unbinding Kinetics through Machine Learning. *J. Chem. Inf. Model.* **2020**, *60* (12), 5946−5956.
(42) Liu, H.; Su, M.; Lin, H.-X.; Wang, R.; Li, Y. Public Data Set of Protein−Ligand Dissociation Kinetic Constants for Quantitative Structure−Kinetics Relationship Studies. *ACS Omega* **2022**, *7* (22), 18985−18996.
(43) Schopf, F. H.; Biebl, M. M.; Buchner, J. The HSP90 chaperone machinery. *Nat. Rev. Mol. Cell Biol.* **2017**, *18* (6), 345−360.
(44) Prodromou, C. Mechanisms of Hsp90 regulation. *Biochem. J.* **2016**, *473* (16), 2439−2452.
(45) Hoter, A.; El-Sabban, M. E.; Naim, H. Y. The HSP90 Family: Structure, Regulation, Function, and Implications in Health and Disease. *Int. J. Mol. Sci.* **2018**, *19* (9), No. 2560.
(46) Sanchez-Martin, C.; Serapian, S. A.; Colombo, G.; Rasola, A. Dynamically Shaping Chaperones. Allosteric Modulators of HSP90 Family as Regulatory Tools of Cell Metabolism in Neoplastic Progression. *Front. Oncol.* **2020**, *10*, 1177.
(47) Birbo, B.; Madu, E. E.; Madu, C. O.; Jain, A.; Lu, Y. Role of HSP90 in Cancer. *Int. J. Mol. Sci.* **2021**, *22* (19), No. 10317, DOI: 10.3390/ijms221910317.
(48) Li, L.; Wang, L.; You, Q. D.; Xu, X. L. Heat Shock Protein 90 Inhibitors: An Update on Achievements, Challenges, and Future Directions. *J. Med. Chem.* **2020**, *63* (5), 1798−1822.
(49) Rowlands, M.; McAndrew, C.; Prodromou, C.; Pearl, L.; Kalusa, A.; Jones, K.; Workman, P.; Aherne, W. Detection of the ATPase activity of the molecular chaperones Hsp90 and Hsp72 using the TranscreenerTM ADP assay kit. *J. Biomol. Screening* **2010**, *15* (3), 279−286.
(50) Southworth, D. R.; Agard, D. A. Species-dependent ensembles of conserved conformational states define the Hsp90 chaperone ATPase cycle. *Mol. Cell* **2008**, *32* (5), 631−640.
(51) Krukenberg, K. A.; Street, T. O.; Lavery, L. A.; Agard, D. A. Conformational dynamics of the molecular chaperone Hsp90. *Q. Rev. Biophys.* **2011**, *44* (2), 229−255.
(52) Henot, F.; Rioual, E.; Favier, A.; Macek, P.; Crublet, E.; Josso, P.; Brutscher, B.; Frech, M.; Gans, P.; Loison, C.; Boisbouvier, J. Visualizing the transiently populated closed-state of human HSP90 ATP binding domain. *Nat. Commun.* **2022**, *13* (1), No. 7601.
(53) Amaral, M.; Kokh, D. B.; Bomke, J.; Wegener, A.; Buchstaller, H. P.; Eggenweiler, H. M.; Matias, P.; Sirrenberg, C.; Wade, R. C.; Frech, M. Protein conformational flexibility modulates kinetics and thermodynamics of drug binding. *Nat. Commun.* **2017**, *8* (1), No. 2276.
(54) Bruce, N. J.; Ganotra, G. K.; Richter, S.; Wade, R. C. KBbox: A Toolbox of Computational Methods for Studying the Kinetics of Molecular Binding. *J. Chem. Inf. Model.* **2019**, *59* (9), 3630−3634.
(55) Lu, H.; Iuliano, J. N.; Tonge, P. J. Structure-kinetic relationships that control the residence time of drug-target complexes: insights from molecular structure and dynamics. *Curr. Opin. Chem. Biol.* **2018**, *44*, 101−109.
(56) Spagnuolo, L. A.; Eltschkner, S.; Yu, W.; Daryaee, F.; Davoodi, S.; Knudson, S. E.; Allen, E. K.; Merino, J.; Pschibul, A.; Moree, B.; et al. Evaluating the Contribution of Transition-State Destabilization to Changes in the Residence Time of Triazole-Based InhA Inhibitors. *J. Am. Chem. Soc.* **2017**, *139* (9), 3417−3429.
(57) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739−1749.
(58) Rajagopal, K.; Sri, V. B.; Byran, G.; Gomathi, S. Pyrazole Substituted 9-Anilinoacridines as HER2 Inhibitors Targeting Breast Cancer - An in-silico approach. *Curr. Drug Res. Rev.* **2022**, *14* (1), 61−72.
(59) Thirunavukkarasu, M. K.; Karuppasamy, R. Drug repurposing combined with MM/PBSA based validation strategies towards MEK inhibitors screening. *J. Biomol. Struct. Dyn.* **2022**, *40* (22), 12392−12403.
(60) Muthu Kumar, T.; Rohini, K.; Nivya, J.; Shanthi, V.; Ramanathan, K. Discovery of potent Covid-19 main protease inhibitors using integrated drug-repurposing strategy. *Biotechnol. Appl. Biochem.* **2021**, *68* (4), 712−725, DOI: 10.1002/bab.2159.
(61) Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* **2019**, *119* (16), 9478−9508.
(62) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12* (85), 2825−2830, DOI: 10.48550/arXiv.1201.0490.
(63) Dixon, S. L.; Jianxin, D.; S, E.; B, C. D. V.; Woody, S.; Matthew, P. R. AutoQSAR: an automated machine learning tool for best-practice quantitative structure-activity relationship modeling. *Future Med. Chem.* **2016**, *8* (15), 1825−1839.
(64) RDKit: Open-source cheminformatics. https://www.rdkit.org.
(65) Dong, J.; Cao, D. S.; Miao, H. Y.; Liu, S.; Deng, B. C.; Yun, Y. H.; Wang, N. N.; Lu, A. P.; Zeng, W. B.; Chen, A. F. ChemDes: an integrated web-based platform for molecular descriptor and fingerprint computation. *J. Cheminf.* **2015**, *7*, 60.
(66) Wang, Z.; Zheng, L.; Liu, Y.; Qu, Y.; Li, Y. Q.; Zhao, M.; Mu, Y.; Li, W. OnionNet-2: A Convolutional Neural Network Model for Predicting Protein-Ligand Binding Affinity Based on Residue-Atom Contacting Shells. *Front. Chem.* **2021**, *9*, No. 753002.
(67) Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28* (28), 31−36.
(68) Agu, P. C.; Afiukwa, C. A.; Orji, O. U.; Ezeh, E. M.; Ofoke, I. H.; Ogbu, C. O.; Ugwuja, E. I.; Aja, P. M. Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. *Sci. Rep.* **2023**, *13* (1), No. 13398.
(69) Pinzi, L.; Rastelli, G. Molecular Docking: Shifting Paradigms in Drug Discovery. *Int. J. Mol. Sci.* **2019**, *20* (18), No. 4331, DOI: 10.3390/ijms20184331.
(70) Stanzione, F.; Giangreco, I.; Cole, J. C. Use of molecular docking computational tools in drug discovery. *Prog. Med. Chem.* **2021**, *60*, 273−343.
(71) Tiwari, P. B.; Wang, X.; He, J.; Darici, Y. Analyzing surface plasmon resonance data: choosing a correct biphasic model for interpretation. *Rev. Sci. Instrum.* **2015**, *86* (3), No. 035001.
(72) Lund-Katz, S.; Nguyen, D.; Dhanasekaran, P.; Kono, M.; Nickel, M.; Saito, H.; Phillips, M. C. Surface plasmon resonance analysis of the mechanism of binding of apoA-I to high density lipoprotein particles. *J. Lipid Res.* **2010**, *51* (3), 606−617.
(73) Bhomia, R.; Trivedi, V.; Coleman, N. J.; Mitchell, J. C. The thermal and storage stability of bovine haemoglobin by ultraviolet-visible and circular dichroism spectroscopies. *J. Pharm. Anal.* **2016**, *6* (4), 242−248.
(74) Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **2006**, *1* (6), 2876−2890.
(75) Böhm, G.; Muhr, R.; Jaenicke, R. Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng.* **1992**, *5* (3), 191−195.
(76) Lammert, H.; Schug, A.; Onuchic, J. N. Robustness and generalization of structure-based models for protein folding and function. *Proteins* **2009**, *77* (4), 881−891.
(77) Noel, J. K.; Levi, M.; Raghunathan, M.; Lammert, H.; Hayes, R. L.; Onuchic, J. N.; Whitford, P. C. SMOG 2: A Versatile Software Package for Generating Structure-Based Models. *PLoS Comput. Biol.* **2016**, *12* (3), No. e1004794.
(78) Clementi, C.; Nymeyer, H.; Onuchic, J. N. Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.* **2000**, *298* (5), 937−953.

(79) Noel, J. K.; Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N. SMOG@ctbp: simplified deployment of structure-based models in GROMACS. *Nucleic Acids Res.* **2010**, *38* (Web Server issue), W657−661.
(80) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185* (2), 604−613.
(81) Bussi, G.; Tribello, G. A. Analyzing and Biasing Simulations with PLUMED. *Methods Mol. Biol.* **2019**, *2022*, 529−578.
(82) Güldenhaupt, J.; Amaral, M.; Kotting, C.; Schartner, J.; Musil, D.; Frech, M.; Gerwert, K. Ligand-Induced Conformational Changes in HSP90 Monitored Time Resolved and Label Free-Towards a Conformational Activity Screening for Drug Discovery. *Angew. Chem., Int. Ed.* **2018**, *57* (31), 9955−9960.
(83) Belkacemi, Z.; Bianciotto, M.; Minoux, H.; Lelievre, T.; Stoltz, G.; Gkeka, P. Autoencoders for dimensionality reduction in molecular dynamics: Collective variable dimension, biasing, and transition states. *J. Chem. Phys.* **2023**, *159* (2), No. 024122, DOI: 10.1063/5.0151053.
(84) Yan, F.; Liu, X.; Zhang, S.; Zhang, Q.; Chen, J. Understanding conformational diversity of heat shock protein 90 (HSP90) and binding features of inhibitors to HSP90 via molecular dynamics simulations. *Chem. Biol. Drug Des.* **2020**, *95* (1), 87−103.
(85) Yang, F.; Wang, Y.; Yan, D.; Liu, Z.; Wei, B.; Chen, J.; He, W. Binding Mechanism of Inhibitors to Heat Shock Protein 90 Investigated by Multiple Independent Molecular Dynamics Simulations and Prediction of Binding Free Energy. *Molecules* **2023**, *28* (12), No. 4792, DOI: 10.3390/molecules28124792.
(86) Chen, J.; Wang, J.; Lai, F.; Wang, W.; Pang, L.; Zhu, W. Dynamics revelation of conformational changes and binding modes of heat shock protein 90 induced by inhibitor associations. *RSC Adv.* **2018**, *8* (45), 25456−25467.
(87) Prajapati, J. D.; Onuchic, J. N.; Sanbonmatsu, K. Y. Exploring the Energy Landscape of Riboswitches Using Collective Variables Based on Tertiary Contacts. *J. Mol. Biol.* **2022**, *434* (18), No. 167788.
(88) García, A. E.; Jose′ N, O. Folding a protein in a computer: An atomic description of the folding/unfolding of protein A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (24), 13898−13903.
(89) Whitford, P. C.; Miyashita, O.; Levy, Y.; Onuchic, J. N. Conformational Transitions of Adenylate Kinase: Switching by Cracking. *J. Mol. Biol.* **2007**, *366* (5), 1661−1671, DOI: 10.1016/j.jmb.2006.11.085.
(90) Levi, M.; Bandarkar, P.; Yang, H.; Wang, A.; Mohanty, U.; Noel, J. K.; Whitford, P. C. Using SMOG 2 to Simulate Complex Biomolecular Assemblies. *Methods Mol. Biol.* **2019**, *2022*, 129−151.
(91) de Oliveira, A. B.; Contessoto, V. G.; Hassan, A.; Byju, S.; Wang, A.; Wang, Y.; Dodero-Rojas, E.; Mohanty, U.; Noel, J. K.; Onuchic, J. N.; Whitford, P. C. SMOG 2 and OpenSMOG: Extending the limits of structure-based models. *Protein Sci.* **2022**, *31* (1), 158−172.
(92) Mascarenhas, N. M.; Gosavi, S. Understanding protein domain-swapping using structure-based models of protein folding. *Prog. Biophys. Mol. Biol.* **2017**, *128*, 113−120.

https://doi.org/10.1021/jacsau.4c00123
JACS Au 2024, 4, 1632−1645