



Chinese Pharmaceutical Association
Institute of Materia Medica, Chinese Academy of Medical Sciences

Acta Pharmaceutica Sinica B

www.elsevier.com/locate/apsb
www.sciencedirect.com



ORIGINAL ARTICLE

Discovery of highly potent phosphodiesterase-1 inhibitors by a combined-structure free energy perturbation approach



Zhe Li ^{a,†}, Mei-Yan Jiang ^{a,†}, Runduo Liu ^{a,†}, Quan Wang ^a,
Qian Zhou ^{a,b}, Yi-You Huang ^{a,b}, Yinuo Wu ^{a,*}, Chang-Guo Zhan ^{c,*},
Hai-Bin Luo ^{a,b,d,*}

^aState Key Laboratory of Anti-Infective Drug Discovery and Development, School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, China

^bKey Laboratory of Tropical Biological Resources of Ministry of Education, School of Pharmaceutical Sciences, Hainan University, Haikou 570228, China

^cMolecular Modeling and Biopharmaceutical Center and Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, Lexington, KY 40536, USA

^dSong Li' Academician Workstation of Hainan University (School of Pharmaceutical Sciences), Sanya 572000, China

Received 4 April 2024; received in revised form 28 May 2024; accepted 18 June 2024

KEY WORDS

Free energy perturbation;
Drug design;
Phosphodiesterase 1;
Relative binding free energy;
Molecular simulation

Abstract Accurate receptor/ligand binding free energy calculations can greatly accelerate drug discovery by identifying highly potent ligands. By simulating the change from one compound structure to another, the relative binding free energy (RBFE) change can be calculated based on the theoretically rigorous free energy perturbation (FEP) method. However, existing FEP-RBFE approaches may face convergence challenges due to difficulties in simulating non-physical intermediate states, which can lead to increased computational costs to obtain the converged results. To fundamentally overcome these issues and accelerate drug discovery, a new combined-structure RBFE (CS-FEP) calculation strategy was proposed, which solved the existing issues by constructing a new alchemical pathway, smoothed the alchemical transformation, increased the phase-space overlap between adjacent states, and thus significantly increased the convergence and accelerated the relative binding free energy calculations. This method was extensively tested in a practical drug discovery effort by targeting phosphodiesterase-1 (PDE1). Starting from a PDE1 inhibitor (compound **9**, $IC_{50} = 16.8 \mu\text{mol/L}$), the CS-FEP guided hit-

*Corresponding authors.

E-mail addresses: wyinuo3@mail.sysu.edu.cn (Yinuo Wu), zhan@uky.edu (Chang-Guo Zhan), hbluo@hainanu.edu.cn (Hai-Bin Luo).

[†]These authors made equal contributions to this work.

Peer review under the responsibility of Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences.

<https://doi.org/10.1016/j.apsb.2024.06.021>

2211-3835 © 2024 Published by Elsevier B.V. on behalf of Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

to-lead optimizations resulted in a promising lead (**11b** and its mesylate salt formulation **11b**-Mesylate, $IC_{50} = 7.0$ nmol/L), with ~ 2400 -fold improved inhibitory activity. Further experimental studies revealed that the lead showed reasonable metabolic stability and significant anti-fibrotic effects *in vivo*.

© 2024 Published by Elsevier B.V. on behalf of Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Binding between biological macromolecules (as drug targets) and small molecules plays important roles in determining the pharmacological effects of drugs. However, accurate calculations of the binding free energies remain one of the greatest challenges in computational chemistry and drug design. Binding free energy prediction methods such as docking scores¹, molecular mechanics/Poisson-Boltzmann surface area (MM/PBSA)^{2,3}, and linear interaction energy (LIE)^{4,5}, have been widely used in drug design, but these methods have problems in accuracy and thus usually are unable to distinguish the binding affinity differences between different small molecules. In recent years, due to the continuous development of computer hardware, using theoretically more rigorous and extensive molecular dynamics (MD) simulations-based statistical mechanics methods to calculate binding free energy have become possible. Among these statistical mechanics methods, the free energy perturbation (FEP) method⁶⁻¹¹ is recognized as a reliable one. Our previously developed restraint energy distribution (RED) function based absolute binding free energy (ABFE)-FEP approach was successfully applied in the virtual screening and discovery of novel and potent inhibitors against various drug targets¹²⁻¹⁵. Especially, during COVID-19, dipyrindamole was discovered against SARS-CoV-2 main protease M^{pro} via our accelerated ABFE-FEP-based large-scale virtual screening from over 12,000 compounds and achieved remarkable clinical treatment effect (the sensitivity of M^{pro} to different assay conditions have also been discussed)^{12,14,16-20}; the related work was highlighted in *Nature Computational Science*²¹.

Further, for the “hit-to-lead” structure optimization stage of drug discovery, FEP-based relative binding free energy (RBFEE) calculation²²⁻²⁵ may be another ideal computational strategy. Conceptually, by computing the work of artificially converting one ligand into another, the FEP method can be used to calculate the difference in the binding free energy with a given receptor between two different ligands, which is also known as an RBFEE calculation. In the early stages of FEP-RBFEE methods using one-step approach (or the “concerted” protocol), the underdeveloped nature of softcore potentials and the simultaneous modification of electrostatic and van der Waals interactions often led to the “endpoint catastrophe”^{26,27} and “particle collapse”²⁸ problems. To alleviate the “endpoint catastrophe”, the three-steps approach (or the “stepwise” protocol) was proposed, which contains at least three perturbation steps (including the decharge, changing vdW parameters, and recharge). With the development of various softcore potentials and in the modern molecular dynamics simulation software (*e.g.*, the ability to decouple the electrostatic and van der Waals interactions asynchronously in NAMD²⁹), the issues of the “endpoint catastrophe” and “particle collapse” have been significantly mitigated. Despite these advancements, other challenges persist in RBFEE calculations. The computational cost

associated with multiple perturbation steps remains a concern, and the lack of guaranteed neutrality in the simulation systems during the alchemical transformation necessitates additional corrections for electrostatic finite-size effects³⁰. Furthermore, when there is a substantial disparity in the van der Waals terms between the initial and final functional groups, the alchemical transformation can engender significant perturbations, potentially leading to non-convergent and low-precision free energy calculations.

In recent years, based on one-step or three-step approaches, researchers have proposed more advanced techniques such as enhanced sampling and softcore potential to improve calculation convergence and accuracy. For instance, Chipot and colleagues introduced a hybrid single-dual-topology free-energy perturbation approach²⁹. Oshima has proposed a new FEP scheme by introducing a modified Hamiltonian. This scheme effectively reduces the computational cost associated with particle mesh Ewald calculations during the FEP computation process³¹. Lee et al. pioneered an alchemical enhanced sampling technique that, in the case of T4-lysozyme, demonstrated superior sampling efficacy compared to the REST2-like enhanced sampling technique³². Furthermore, Lee et al. introduced a novel variant of soft-core potentials known as the smoothstep softcore (SSC) potentials^{33,34}. Although these efforts have improved the RBFEE calculation, the inherent limitations coming from the design of perturbation pathways still exist for both one-step and three-step approaches. For example, the simultaneous change of Coulomb and vdW parameters in one-step approach will contribute to the instability of the free energy predictions. The three-step approach that completely decoupled the change of Coulomb and vdW parameters is not smooth enough in the perturbation pathways, resulting in lower efficiency. NAMD has designed a special vdW and electrostatic coupling parameter evolution scheme that utilizes a single λ parameter to control the asynchronous changes in vdW and electrostatic interactions²⁹. This approach overcomes the “endpoint catastrophe” and “particle collapse” problems associated with traditional methods. However, it still cannot ensure that the system maintains charge neutrality throughout the simulation process.

To fundamentally solve these problems and bring further development to the FEP-based methods, a new, more robust combined state alchemical transformation approach (abbreviated as CS-FEP below) is proposed in this study (as briefly depicted in Fig. 1) to transform the ligand smoothly and efficiently from the reference structure to the target structure while avoiding simultaneous transformation of both the vdW and Coulomb terms for each single atom. At the same time, the CS-FEP approach exhibited a higher degree of phase-space overlap in the vdW transformation than the traditional approaches. This new approach was extensively tested by using an external test set containing eight drug targets, about 200 ligands, and more than 350 pairs of RBFEE calculations. For each pair of RBFEE calculations, multiple

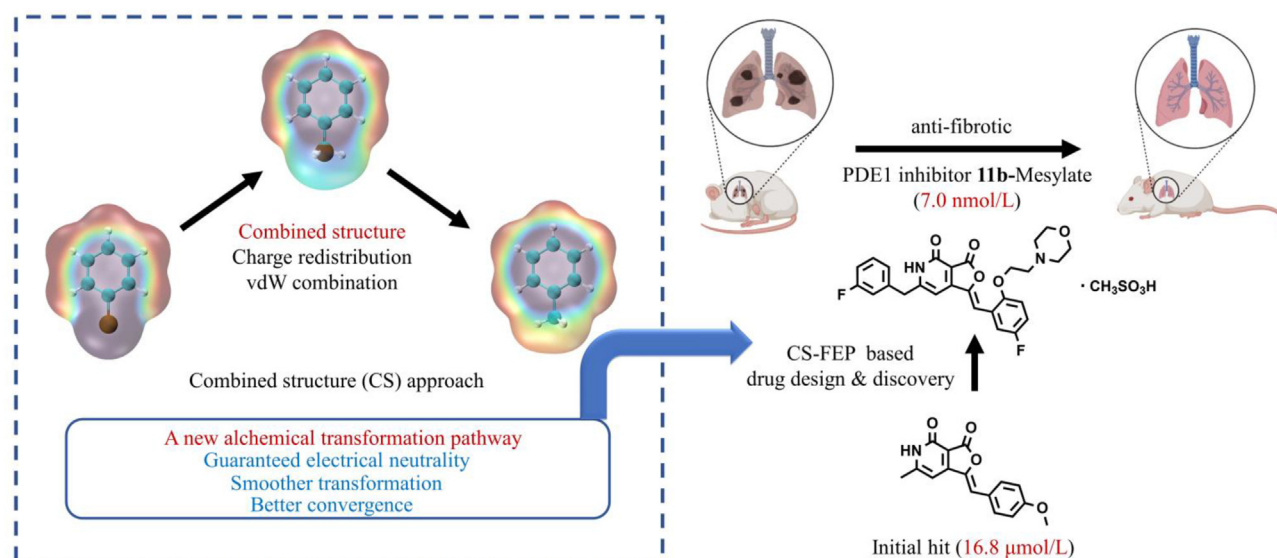


Figure 1 Overview of the CS-FEP approach which was validated by the hit-to-lead discovery of PDE1 inhibitors with anti-fibrotic activities. The molecular surface electrostatic potential map was generated by VMD³⁷ in combination with the Multiwfn program³⁸.

alchemical pathways were calculated, and the convergence behavior was extensively analyzed. With the smooth transformation, the convergence of RBFE calculations is greatly improved. The performance between the three-steps method and the CS-FEP method was directly compared based on a CDK2 system consisting of 25 pairs, and results indicated that the CS-FEP approach obtained better results while requiring less computational cost. For the minor changes in structures, such as the 11 ligands of thrombin, remarkable accuracy (with $R^2 > 0.65$) can be achieved by even using a single window (associated with only one λ value) for each alchemical transformation.

The developed CS-FEP protocol has been used successfully in a practical drug design effort to design and discover potent phosphodiesterase-1 (PDE1) inhibitors with a novel scaffold and anti-fibrotic effects *in vivo*. Our previous studies^{35,36} identified PDE1 as a potentially important target for the pulmonary fibrosis (PF) preventative treatment. Herein, to rapidly discover potent PDE1 inhibitors as anti-fibrotic agents, a virtual screening was performed to discover a “hit” with a novel scaffold and weak inhibitory activity (compound **9**, $IC_{50} = 16.8 \mu\text{mol/L}$). After the CS-FEP-guided hit-to-lead structural optimizations, we successfully predicted, synthesized, and discovered a lead compound (**11b**, and its mesylate salt formulation **11b-Mesylate**) with an IC_{50} value of 7.0 nmol/L, desirable liver microsomal stability, and remarkable anti-fibrotic effects *in vivo*. We anticipate that the robust CS-FEP approach developed in this study will be used widely in various practical rational drug design and discovery efforts, particularly in the hit-to-lead and lead optimizations.

2. Result and discussion

2.1. The advantages of the CS-FEP approach

Unlike the usually used one-step and three-step RBFE protocols, in the CS-FEP approach, a new “combined state” is added to the thermodynamic cycle (Fig. 2A), which helps to increase both the speed and the phase-space overlap during the alchemical transformation. Additionally, the CS-FEP approach can be used to

overcome the aforementioned computational problems existing in both the one-step and three-step protocols. According to the thermodynamic cycle depicted in Fig. 2A, RBFE for a specific pair of compounds, A and B, is calculated using Eq. (1):

$$\Delta\Delta G = \Delta G_B - \Delta G_A = \Delta G_{\text{CS to A}}^{\text{lig}} - \Delta G_{\text{CS to B}}^{\text{lig}} + \Delta G_{\text{CS to B}}^{\text{com}} - \Delta G_{\text{CS to A}}^{\text{com}} \quad (1)$$

Based on the structures of ligands A (reference structure) and B (target structure), the combined structure contains their common atoms, as well as the unique atoms from both ligands A and B (the interactions between unique atoms from different ligands are excluded). In the combined structure, the unique atoms are not charged, and the charges of the common atoms are calculated as described in the “Practical RBFE calculations using the CS-FEP method” subsection (Supporting Information Section S1) to ensure the smooth transformation. As explicitly stated by Loeffler et al. (Supporting Information Section S3 of their research)³⁹, most MD codes do not support maintaining charge neutrality within each λ step, meaning that unless charges are turned off for all atoms, the overall system’s charge will change with λ variation. Especially for the two “decharge” steps in the three-step method, the process of alchemically changing charges cannot guarantee that the system remains charge-neutral throughout the λ -changing process, even if the charges in the initial and final states are consistent. Therefore, our CS-FEP strategy involves redistributing the charges of common atoms on the combined structures (detailed in Section S1). This ensures that during the linear change of λ , the net charge of the entire molecule remains 0. The transformation from the reference structure to the combined structure includes three types of changes. For the common atoms, both the charge parameters and vdW parameters may change, depending on the specific atom types involved. In cases where common atoms have different atom types and vdW parameters, our approach simultaneously modifies both the vdW parameters and charges of these atoms during the alchemical simulations. For the unique atoms in reference structure, only charge parameters will disappear. For the unique atoms in the target structure, only vdW parameters will appear. The transformation from the combined

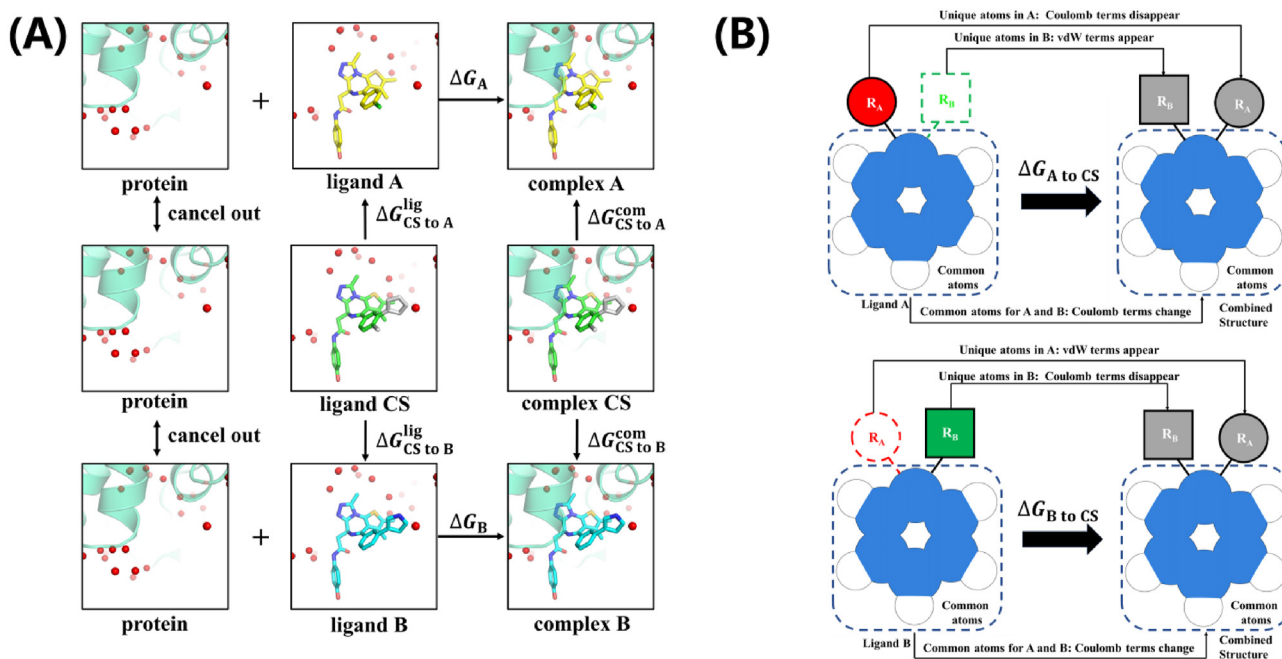


Figure 2 (A) The thermodynamic cycle used in the combined structure approach. The light green cartoons represent the protein while the red spheres represent the water molecules. The yellow and cyan sticks represent the reference ligand (ligand A) and the target ligand (ligand B), respectively. For the CS molecules, the unique region with only vdW terms are shown in gray sticks, and the common region is represented by green sticks. (B) Illustration of alchemical transformations from ligand A or ligand B to the combined structure of ligands A and B. R_A and R_B represent the respective unique substituents of ligands A and B. The groups with gray color represent the substituents with only the vdW terms. The blue dotted box represents the region of the common atoms of ligands A and B.

structure to the target structure can be understood in the same way. Briefly speaking, both the charge and vdW parameters will change during the transformation, but only either the vdW or the charge parameters will change for each part of the structure (Fig. 2B).

The CS-FEP approach offers numerous advantages, such as maintaining system neutrality throughout the entire alchemical transformation process, avoiding the “endpoint catastrophe” and “particle collapse” issues through the asynchronous transformation of the vdW and electrostatic terms of the same atoms and providing a smoother alchemical transformation pathway for the efficient RBE calculations.

2.2. The CS-FEP method exhibits a high degree of phase-space overlap in the vdW transformation

For a primary reason concerning why the combined structure approach has better phase-space overlap than the three steps approach, the vdW terms do not change uniformly with λ . As a result, within the three-step approach, in changing the vdW term of the unique atoms, there exist multiple intermediate states with their steric sizes much smaller than both the initial and the target state sizes. The substantial change in the steric size will further have a larger influence on the atomic positions in phase space. Compared with the three-step approach in which the steric size first decreases and then increases substantially, only the unique atoms will contribute to the change from the combined structure to the initial/target structure in the combined structure approach. As a result, the combined structure approach is expected to have a much smoother transformation.

For validation of this concept, as shown in Fig. 3, we conducted the alchemical transformations for the vdW term of the

unique atoms of 31 to 30 (CDK2 molecule pair), with the charge parameters of all atoms set to 0. For the three-step approach, the ΔG changes indicate that the overall steric size first decreased (λ from 0 to 0.5) and then increased (λ from 0.7 to 1.0). In addition, there are other indicators, all showing that the phase-space overlap is better for the combined structure approach. First, in the three-step approach, the $\delta\Delta G/\delta\lambda$ values differ greatly between different λ windows, indicating the large phase-space differences between these λ windows; in comparison, the $\delta\Delta G/\delta\lambda$ values are closer to each other between different λ windows in the combined structure approach (in Fig. 3B). Second, the “overlap degree” values (that range from 0 to 0.5; the larger the value, the better the phase-space overlap) proposed by Bennett⁴⁰ were calculated for all windows (as shown in Fig. 3A), showing the larger overlap values for the combined structure approach. Additionally, the agreement between the original simulation and the reweighted results from adjacent windows can also indicate the phase-space overlaps between these windows. Thus, ΔG values were also calculated from the reweighted probability distributions from adjacent windows. The calculation method for the “overlap degree” values and the reweighting heat map is given in Supporting Information Section S10. As indicated explicitly in Fig. 3C for the heat map showing the difference between the original and reweighted ΔG values, compared to the three-step approach, the combined structure approach showed better agreement for the ΔG values between the original and reweighted results.

2.3. Superior performance of the CS-FEP method

The CS-FEP method was tested extensively by using a test set containing 8 targets and more than 150 ligands, the same dataset

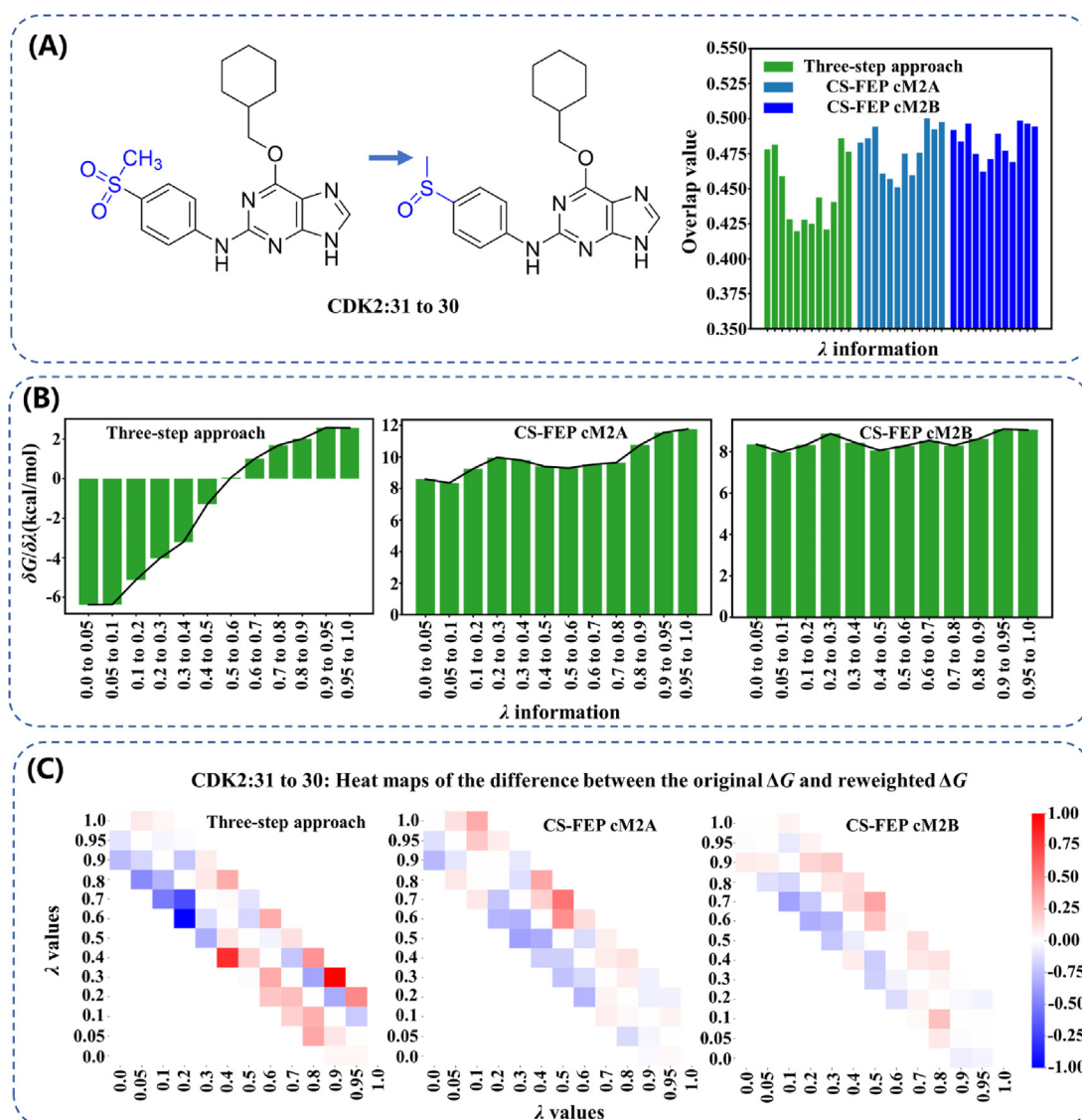


Figure 3 Calculated results of the perturbation of vdW term in the molecule pair 31 to 30 using the three-step and CS-FEP approaches. (A) The overlap degree between adjacent windows obtained from the two approaches; the λ information for each group of bar charts is consistent with that in panel B and has been omitted. (B) The histograms of $\delta G / \delta \lambda$ obtained from both approaches. (C) Heat maps of the difference between the original and reweighted ΔG (with the unit of kcal/mol). The terms “cM2A” and “cM2B” in the figure represent the processes of converting the combined structure in complex to the reference molecule and the target molecule, respectively.

used previously in Wang’s study²² for comparison. More than 300 pairs of RBF calculations were performed for the benchmark test. The cycle closure program wcc⁴¹ was further used to obtain the ΔG values. The correlation between the CS-FEP predicted and experimental values of ΔG for the eight systems are shown in Fig. 4A. There are only 35 molecules (17%) whose predicted binding free energies deviated from the experimentally determined binding free energies by >1.0 kcal/mol. As shown in Table 1, the R^2 values of CS-FEP predicted results and experimental results are greater than 0.5 for most of the systems. All the mean absolute error (MAE) and root-mean-square error (RMSE) values of eight tested systems are less than 1.0 kcal/mol, demonstrating the high accuracy of our CS-FEP approach.

The results from currently widely used FEP+²² and Amber TI²³ were also listed for comparison, as given in Table 1 and Supporting Information Table S1 (in Supporting Information Section S7). For the FEP+, although REST enhanced sampling method was used to improve calculation accuracy, only 2 out of the 8 systems showed better MAE and RMSE than CS-FEP, and only 2 out of the 8 systems showed better R^2 than CS-FEP, indicating the results obtained from CS-FEP are at least comparable to the FEP+, if not better. For each of the indicators (including MAE, RMSE and R^2), CS-FEP outperforms the Amber TI method.

To stringently compare the CS-FEP approach with the currently acknowledged premier alchemical transformation protocol for FEP-RBFE, *i.e.*, the three-step approach, we conducted a

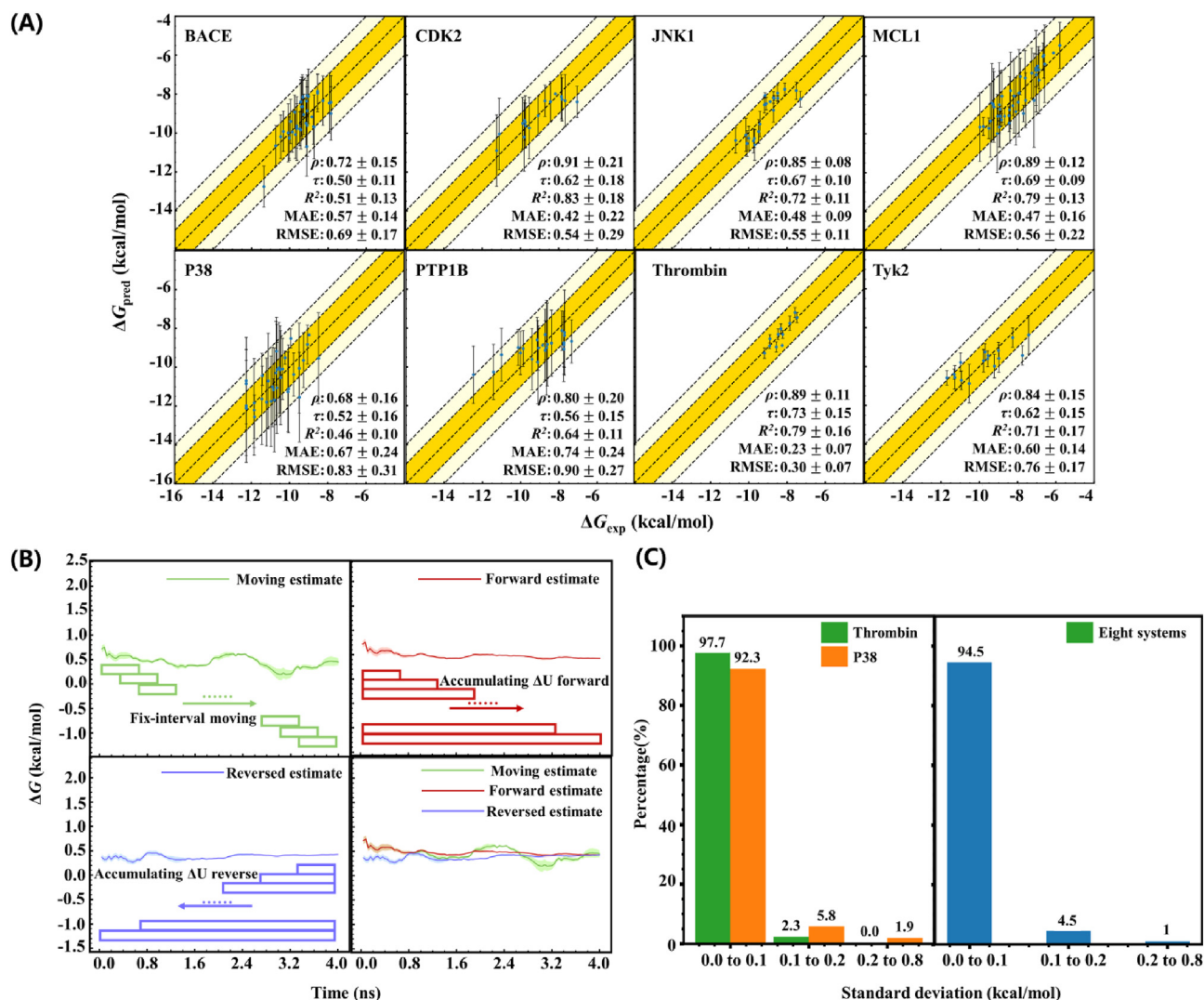


Figure 4 (A) Correlations between the CS-FEP-predicted and experimental ΔG values for the eight tested targets. The scattered points appearing in the deep yellow area indicate that the differences between the experimental values and the calculated values are within ± 1 kcal/mol. (B) Time convergence analysis. The green, red and blue lines represent the moving, forward and reverse estimate, respectively. (C) The standard deviations of ΔG between the adjacent windows involved in the thrombin, P38 and the whole of eight systems, which were calculated by the time-moving analysis.

Table 1 Summary of the Mean Absolute Error (MAE), Root-mean-square Error (RMSE), and R^2 for the ΔG predicted by FEP+, AMBER-TI, and CS-FEP after cycle closure. The best MAE, RMSE, and R^2 for each target were marked by bold.

System	FEP+			AMBER-TI			CS-FEP		
	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2
Bace	0.67	0.85	0.61	1.03	1.32	0.19	0.57	0.69	0.51
CDK2	0.88	1.04	0.23	0.90	1.08	0.22	0.42	0.54	0.83
Jnk1	1.06	1.14	0.72	0.90	1.13	0.22	0.48	0.55	0.72
MCL1	0.84	1.04	0.60	1.24	1.48	0.42	0.47	0.56	0.79
P38	0.86	0.97	0.43	1.28	1.62	0.15	0.67	0.83	0.46
Ptp1b	0.61	0.80	0.64	0.76	1.01	0.50	0.74	0.90	0.64
Thrombin	0.42	0.54	0.50	0.37	0.51	0.57	0.23	0.30	0.79
Tyk2	0.45	0.57	0.79	0.89	1.13	0.33	0.60	0.76	0.71
overall	0.76	0.93	0.66	1.01	1.30	0.44	0.55	0.69	0.74

comparative analysis of their respective outcomes while maintaining uniformity in simulation software, force fields, initial conformations, molecular pair mappings and sampling algorithms. More computational details are provided in [Supporting Information Section S2](#). The CDK2 system consisting of 25 ligand pairs was used for the comparison. For each alchemical transformation in both three-steps and the CS-FEP protocol, we employed similarly the same 13 λ windows, with 4-ns of simulation in each window. Since the thermodynamic cycle of the three-step method involves six alchemical transformations, there are 26 additional λ -windows in the three-steps approach compared to the CS-FEP approach. As evidenced by [Table 2](#), despite using fewer λ windows, the CS-FEP method had better RMSE and MUE performance for both cycle-closure $\Delta\Delta G$ and ΔG than the three-step method. In terms of linear correlation statistical analysis, the CS-FEP calculations (τ of 0.62 ± 0.18) also exhibit a marginally superior performance compared to the three-step method (τ of 0.56 ± 0.17). More detailed results are given in the “cdk2_three_steps_protocol_data” sheet in the [Supporting Information](#). The consistency between the CS-FEP and the classical three-step method is discussed in [Supporting Information Section S4](#) and [Fig. S1](#).

2.4. Standard deviations are below 0.1 kcal/mol for the vast majority of λ windows

In the CS-FEP convergence analysis, we first carried out three analyses related to the time convergence ([Fig. 4B](#)). A standard way to check the time convergence of ΔG is the time-forward estimate. In the time-forward estimate, the ΔU data (which are the differences in internal energy between adjacent states) required for the free energy calculation was accumulated by starting from the first frame to the last frame and ΔG was calculated continuously to obtain the curve of ΔG versus the simulation time. The second time-convergence analysis is the time-reverse estimate, which is similar to the time-forward estimate. For the only difference compared to the time-forward estimate, the last frame is used as the starting point with the ΔU data accumulated backward for the ΔG calculation in the time-reverse estimate. Another analysis related to the time convergence is the time-moving estimate analysis. Similar to the time-forward estimate, the time-moving estimate traverses forward all conformations in the trajectory in chronological order of simulation (the detail results of the time convergence analysis could be found in [Supporting Information Section S8](#) and [Fig. S2](#)). For the difference, in the time-moving estimate, the ΔU data used to calculate

ΔG are limited to a fixed-width time interval (which is 0.4 ns in our study) each time, which allows this analysis to capture the transitions of the system at different metastable states.

Following the standard CS-FEP protocol, the RBFEE calculations of the eight systems in this study resulted in more than 20,000 λ windows, and we calculated the standard deviation of ΔG of all these λ windows as the way described above. As shown in [Fig. 4C](#), most of the standard deviations of ΔG between the adjacent λ windows are smaller than 0.1 kcal/mol, accounting for 94.5%, and only 4.5% of them fall in the interval from 0.1 to 0.2 kcal/mol. Although 1% of the standard deviations of ΔG are greater than 0.2 kcal/mol, they are all smaller than 0.8 kcal/mol. These results suggest that the CS-FEP method has good convergence in the calculation of each RBFEE. Further, we selected two representative systems with the good and poor R^2 (thrombin with R^2 of 0.79 ± 0.16 and P38 with R^2 of 0.46 ± 0.10) and plotted the distribution histograms of the standard deviations. As expected, for the λ windows of the thrombin system, the fraction of the standard deviations lying in the range of 0–0.1 kcal/mol (97.7%) was greater than that of P38 (92.3%), with all the standard deviations being within 0.2 kcal/mol. Correspondingly, the proportion of λ windows with larger standard deviations was higher for CDK2 system, further demonstrating the connection between the standard deviation distribution of λ windows and the accuracy of RBFEE predictions. Thus, obtaining the standard deviation based on time-moving analysis can provide a powerful tool for analyzing the convergence of RBFEE calculations. The results obtained also indicated the good convergence of the CS-FEP method. More convergence analysis and conformation analysis are given in [Supporting Information Sections S8, S9](#) and [Fig. S3](#).

2.5. The CS-FEP approach can obtain reasonable results with few λ windows

Since the CS-FEP approach showed remarkable convergence, it is possible to make further attempts to reduce the computational cost. The number of λ windows was reduced to compare the outcomes. We used 13 windows, 6 windows, 3 windows, and even one window for each transformation in the thermodynamic cycle to estimate the $\Delta\Delta G$. The R^2 corresponding to various schemes are given in [Table 3](#). First, we can see from [Table 3](#) that the R^2 of eight systems decreases as the number of windows used decreases, which is consistent with the fact that the larger the overlap of adjacent windows, the higher the computational accuracy. Furthermore, for four (Jnk1, Mcl1, Ptp1b and thrombin) of the eight systems, as the number of windows decreases, the declines

Table 2 Summary of MAE, RMSE, R^2 , and Kendall's τ Coefficient (τ) for CDK2 system using CS-FEP and three-step approaches.

Dataset for statistical analysis	MAE (kcal/mol)	RMSE (kcal/mol)	R^2	τ
$\Delta\Delta G_{\text{CS-FEP}}$	0.49 ± 0.15	0.63 ± 0.20	—	—
$\Delta\Delta G_{\text{one-step}}$	0.97	1.13	—	—
$\Delta\Delta G_{\text{three-step}}$	1.17 ± 0.25	1.40 ± 0.31	—	—
$\Delta G_{\text{CS-FEP}}$	0.42 ± 0.22	0.54 ± 0.29	0.83 ± 0.18	0.62 ± 0.18
$\Delta G_{\text{one-step}}$	0.90	1.08	0.22	0.25
$\Delta G_{\text{three-step}}$	0.82 ± 0.32	1.07 ± 0.45	0.52 ± 0.19	0.56 ± 0.17

Note: Since the correlation statistics for $\Delta\Delta G$ will be significantly affected by the different definitions of the directions of the molecular pairs⁴², thus we only report the RMSE and MAE of the $\Delta\Delta G$. The standard deviations of statistical metrics are calculated by bootstrapping analysis. The calculation results of the one-step method come from the research of Amber-TI²³.

of their R^2 values are minor, which indicates that transformations are quite smooth with good overlap between the neighboring states and, thus, one may potentially use even fewer windows to obtain sufficiently reliable results. At the same time, the R^2 of Tyk2 decreases significantly with the decreasing number of λ windows. The phenomenon may be attributed to the cause that there are many transformations between molecular pairs involving the changes from ring substituents to chain substituents. Although some of the systems such as Tyk2 may be more sensitive to the number of λ windows, in general, the CS-FEP can get reasonable results for most of the systems when even one λ window was used for each transformation, as seen in Table 3.

Besides the number of λ windows, we also tried to decrease the simulation time for the free energy calculations. Compared to the $\Delta\Delta G$ predicted based on the last 1-ns data of each simulation window (as mentioned in Experimental section), we further decreased the simulation time to the last 80 ps of each simulation window. The corresponding R^2 value of each system is shown in Table 3. Even with such a short period of time (the last 80 ps), the R^2 of JNK1, MCL1, thrombin, TYK2 and CDK2 systems are comparable to that obtained by the last 1 ns of simulation data.

2.6. Discovery of novel PDE1 inhibitors with anti-fibrotic effects via the CS-FEP strategy

PF is a chronic, progressive, and devastating lung disease that is characterized by irreversible fibrosis, a decline in lung function and progressive respiratory failure. With the recent rampage of COVID-19 in recent years, several studies^{43–47} have indicated that patients with PF often have a higher risk of developing COVID-19-related complications and a poorer prognosis for COVID-19.

PDE1, belonging to a superfamily of enzymes (PDEs, including 11 PDE gene families), can catalyze hydrolysis of the important intracellular second messengers cAMP and cGMP. Our most recent studies demonstrated that PDE1 might work as a potential drug target for PF^{35,36}. However, there have been few PDE1 inhibitors reported so far. It is highly desired for development of a novel PF preventative treatment to design and discover potent PDE1 inhibitors, especially those with novel scaffolds. To discover novel PDE1 inhibitors as anti-fibrotic agents, we carried out a drug discovery process including virtual screening, CS-FEP guided structure optimization, chemical synthesis, metabolic stability test, and anti-fibrotic effect validation *in vivo*. Virtual screening against the ChemDiv compound library was conducted using an MM-PBSA-based virtual screening pipeline. This approach has been extensively employed in our previous virtual

screening studies^{48,49}. A PDE1 inhibitor **9** (ChemDiv compound ID: 8005-2367) with a novel scaffold, furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione, was discovered as a hit showing weak inhibitory activity (IC₅₀: 16.8 μ M) against PDE1.

Starting from hit **9**, with the assistance of CS-FEP relative binding free energy calculation, a series of furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione derivatives were designed and predicted. Fourteen compounds with both appropriate binding modes and binding free energies were synthesized and tested using previously reported bioassays^{50–53}. The results are summarized in Table 4. Information about the molecular pair map for the RBFE calculations of the PDE1 system can be found in the Supporting Information, specifically in Section S11. The experimental binding affinities converted from IC₅₀ values fit well with the CS-FEP prediction results. To start the structural optimization, we found that compound **9** was anchored in the catalytic pocket through hydrogen-bonding interactions with Gln421, π - π stacking with Phe424, and hydrogen-bonding interactions with Tyk222 mediated by a conserved water. Its benzyldiene group stretched into the metal-binding site but does not sufficiently interact with surrounding amino acids. Our modification started with changing the substituents at the phenyl groups (R₁ and R₂). First, we designed 25 compounds for the initial round of CS-FEP calculations (see the "PDE1_design_cal_compound" sheet in the Supporting Information). Considering a balance of stronger predicted affinity and higher ligand efficiency, we selected five compounds for synthesis and activity testing: **9a**, **9b**, **9c**, **9d**, and **9e**. Compared to compound **9** with 4-methoxyl at the R₁ position, compound **9a** with 3-methoxyl gave an improved IC₅₀ value of 703 nmol/L against PDE1. When changing the substituents to the ortho position, the inhibition was greatly improved. Compound **9b** with a 2-ethoxyl group, compound **9c** with 2-isopropyl group, and compound **9d** with a 2-cyclopropylmethoxy group gave comparable inhibitory activities, with the IC₅₀ values of 147, 216 and 190 nmol/L, respectively. When changing the 2-ethoxyl group to 2-chloro, the inhibitory activity decreased to 614 nmol/L. The predicted binding modes of **9b–e** showed that substituents at the 2-position of phenyl ring stretched into a hydrophobic pocket composing of Phe392, Thr271, and Leu409. Introducing hydrophobic groups with suitable volume might form hydrophobic interaction with the pocket and enhance the inhibition.

In the second round of CS-FEP-guided activity optimization, based on compound **9c**, we designed six compounds with two substitutions on the phenyl ring (see the "PDE1_design_cal_compound" sheet in Supporting Information). Except for the compound with a Cl substitution at position 4, which was

Table 3 Eight systems' R^2 calculated using different stratification strategies and simulation times.

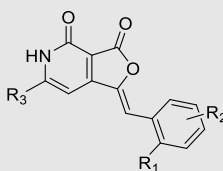
Number of windows	Time for FE calculation	Bace	CDK2	Jnk1	MCL1	P38	Ptp1b	Thrombin	Tyk2
13 ^a	1 ns	0.51	0.83	0.72	0.79	0.46	0.63	0.79	0.71
13	80 ps	0.39	0.67	0.73	0.65	0.22	0.19	0.85	0.74
6 ^b	1 ns	0.45	0.65	0.72	0.79	0.37	0.77	0.72	0.77
3 ^c	1 ns	0.34	0.72	0.68	0.60	0.33	0.61	0.87	0.13
1 ^d	1 ns	0.23	0.73	0.57	0.42	0.33	0.51	0.73	0.10


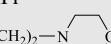
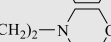
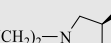
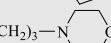
^aThe λ values of 13 windows used for each alchemical transformation are 0.00, 0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 0.95 and 1.00, respectively.

^bThe λ values of 6 windows used for each alchemical transformation are 0.05, 0.20, 0.40, 0.60, 0.80 and 0.95, respectively.

^cThe λ values of 3 windows used for each alchemical transformation are 0.10, 0.50 and 0.90, respectively.

^dThe λ value of this window for each alchemical transformation is 0.50.

Table 4 Molecular structures, inhibitory activity, the experimental and predicted free energies of novel PDE1 inhibitors (the unit of free energy is kcal/mol).


Compd.	R ₁	R ₂	R ₃	IC ₅₀ (nmol/L)	ΔG _{exp} (kcal/mol)	ΔG _{pred} (kcal/mol)
9	H	4-OMe	Me	16,800 ± 3600	-6.51 ± 0.12	-6.51 ± 0.00
9a	H	3-OMe	Me	703 ± 58	-8.39 ± 0.05	-7.74 ± 0.25
9b	OEt	H	Me	147 ± 4	-9.31 ± 0.01	-8.21 ± 0.41
9c	OiPr	H	Me	216 ± 10	-9.08 ± 0.02	-8.55 ± 0.47
9d	OCH ₂ - 	H	Me	190 ± 12	-9.16 ± 0.04	-8.75 ± 0.42
9e	Cl	H	Me	614 ± 69	-8.47 ± 0.07	-7.90 ± 0.49
10a	OiPr	5-F	Me	16 ± 1	-10.62 ± 0.03	-9.35 ± 0.47
10b	OiPr	4-F	Me	709 ± 32	-8.38 ± 0.03	-8.83 ± 0.50
10c	OiPr	5-Br	Me	121 ± 11	-9.43 ± 0.05	-9.13 ± 0.50
10d	OiPr	4-Br	Me	1012 ± 52	-8.17 ± 0.03	-8.47 ± 0.53
10e	OiPr	5-Cl	Me	45 ± 3	-10.01 ± 0.04	-9.25 ± 0.50
11a	O(CH ₂) ₂ - 	5-F	Me	31 ± 3	-10.23 ± 0.05	-10.30 ± 0.48
11b	O(CH ₂) ₂ - 	5-F	3-F-Bn	7.0 ± 0.7	-11.11 ± 0.05	-14.13 ± 0.52
11c	O(CH ₂) ₂ - 	5-F	3-F-Bn	3.8 ± 1.1	-11.48 ± 0.15	-15.07 ± 0.82
11d	O(CH ₂) ₃ - 	5-F	3-F-Bn	8.3 ± 1.2	-11.01 ± 0.08	-14.59 ± 0.82

predicted to have significantly lower affinity than compound **9c**, the other five compounds demonstrated comparable or even better binding affinities under the CS-FEP algorithm's predictions. Therefore, in this design phase, we selected five compounds, namely **10a**, **10b**, **10c**, **10d**, and **10e**, for synthesis and subsequent activity assays. Compared to compound **9c**, **10a** with an additional fluorine atom at 5-position gave the IC₅₀ of 16 nmol/L while **10b** with a fluorine atom at 4-position gave the IC₅₀ of 709 nmol/L

against PDE1. This phenomenon also happened on compound **10c** with 5-Br and compound **10d** with 4-Br, their IC₅₀ values were 121 and 1012 nmol/L, respectively. According to the binding modes of these compounds (Fig. 5), the substituents at the 4-position of phenyl might collide with a water in the metal region, which is unfavorable for binding with the protein. Substituents at the 5-position of phenyl do not have these interactions and thus result in an improved inhibitory activity. However, the

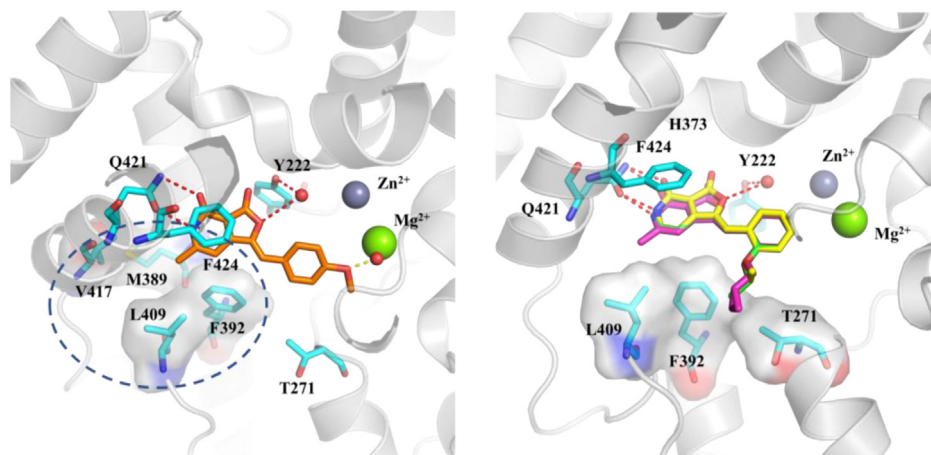


Figure 5 The predicted binding poses of the furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione derivative compounds binding with PDE1 protein. The key residues are shown in cyan sticks. The compounds **9**, **9b**, **9c**, and **9d** are represented by orange, yellow, green, and purple sticks, respectively. The red spheres represent two key waters in the complex. The zinc and magnesium ions are represented by grey and green spheres, respectively. The hydrogen bonds between ligands and protein are shown by red dashes, while the potential clash between the 4-position substituent and the coordinated water are expressed by the yellow dash. The blue dotted circle is the Q2 hydrophobic pocket.

pocket around the 5-position is small and cannot accommodate with bulky substitutions. Thus, compound **10a** with a fluoro atom gave better inhibitory activity than compound **10c** with a bromo atom and compound **10e** with a chloro atom. A flexible ethyl-morpholine chain was used to replace the isopropoxy of compound **10a**, resulting in compound **11a**. The IC_{50} of **11a** against PDE1 was decreased to 31 nmol/L. The binding mode of compound **11a** showed that a Q2 hydrophobic pocket existed nearby the methyl group. Thus, a 3-*F*-benzyl group was introduced at compound **11a**, resulting in compound **11b**, its IC_{50} was improved to 7.0 nmol/L. We also synthesized compounds **11c** and **11d**. Both showed excellent inhibitory activities against PDE1.

To validate the reliability of the predicted binding mode, we selected compound **10a** with the highest ligand efficiency ($LE = 0.44$) and a reasonably potent activity ($IC_{50} = 16$ nmol/L) as a representative PDE1 inhibitor. Based on the originally predicted binding mode, we experimentally introduced mutations in key amino-acid residues Q421, F392, and Y228 to alanine. Meanwhile, to strengthen the reliability of our findings, we further conducted FEP-RBFE calculations on the corresponding changes of amino acids. The computational results showed that changing these key amino acids to alanine would decrease the binding affinity between compound **10a** and PDE1 by 0.5–2 kcal/mol (corresponding to an increase in IC_{50} values by approximately 2.5–30 times). The experimental results showed the IC_{50} of **10a** decreased from 16 nmol/L (wild type) to 127 nmol/L (Q421A), 231 nmol/L (F392A), and 43 nmol/L (Y228A), showing 8-fold, 14-fold, and ~3-fold loss of binding affinities, respectively. The computationally predicted and experimentally observed mutation-induced changes in the binding affinity between compound **10a** and PDE1 are summarized in Table 5. The excellent agreement between the theoretical and experimental data shown in Table 5 provide additional support for the accuracy of the computationally predicted binding modes.

The detailed results of PDE1 are given in Supporting Information Section S11 and Table S2, showing the consistency of accuracy between PDE1 and tested datasets. The RMSE for $\Delta\Delta G$ results is 0.83 kcal/mol for the PDE1 dataset, which is quite close to the RMSE values of the tested systems (0.85 for BACE, 0.94 for P38, 0.84 for TYK2, respectively; please refer to Supporting Information Table S1 in Section S7 for details). The hit-to-lead structural optimizations of PDE1 ligands involved two rounds: the prediction of $\Delta\Delta G$ results within the first round (with RMSE of 0.72 kcal/mol) and the second round (with RMSE of 0.47 kcal/mol) remain reasonably accurate. In the calculation map (Supporting Information Fig. S4), there is only one connection (**11a**–**11b**) between the first round and the second round, and the large perturbation between **11a** and **11b** resulted in a large absolute error of 2.16 kcal/mol of this pair. This is the only pair with the calculation error larger than 2 kcal/mol but contributed substantially

to the ΔG results after cycle closure, resulting in the RMSE of 1.6 kcal/mol of ΔG results. However, using the RMSE of ΔG results after cycle closure instead of the original $\Delta\Delta G$ results cannot directly reflect the accuracy of the RBFE method. The consistency of the accuracy in $\Delta\Delta G$ calculations between PDE1 and tested datasets indicates the reliability of the CS-FEP method.

The metabolic stability of the potent compounds **11a**–**11c** were evaluated by rat liver microsomal experiments *in vitro* (see details in Supporting Information Section S12). Compounds **11a**, **11b**, and **11c** had $t_{1/2}$ (half time) values of 8 min, 29.8 min, and 13.5 min, respectively (Supporting Information Table S3). Since compound **11b** was significantly more stable than **11c**, we chose compound **11b** in its mesylate salt formulation (denoted as **11b**-Mesylate) to be subjected to the bleomycin (BLM)-induced IPF rat model to evaluate its anti-fibrotic effects.

The histopathological analysis of lung tissues in each group was evaluated by means of H&E staining. Compared with the control group, the alveolar septa of BLM-induced rats thickened, and a large number of inflammatory cells infiltrated. After administration of **11b**-Mesylate, the infiltration of inflammatory cells in lung tissue decreased and the alveolar structure improved (Fig. 6A). We also evaluated the deposition of collagen using Masson's trichrome staining and IHC stain for collagen (Fig. 6B–D), which demonstrated that **11b**-Mesylate reduced BLM-induced collagen deposition in the extracellular matrix. Meanwhile, the level of vimentin and α -smooth muscle actin protein (α -SMA) were analyzed by Western blotting method. In the lung tissue of BLM induced rats, the expression of vimentin and α -SMA were increased. Treatment with **11b**-Mesylate decreased the level of α -SMA and vimentin (Fig. 6E). All these results demonstrated that PDE1 inhibitor **11b**-Mesylate ameliorated BLM-induced pulmonary fibrosis.

2.7. Code availability

The code for the free energy calculation is available on GitHub at the following link: https://github.com/zlisysu/CS-FEP_run.

3. Conclusions

In this study, to fundamentally solve the issues of the widely used one-step and three-step FEP approaches for rational drug design, a new combined-structure FEP approach was developed and validated. Particularly, a combined structure will be constructed for the alchemical transformation in the CS-FEP approach, allowing parameters of all atoms of the ligand change simultaneously, and each atom will change either vdW or electrostatic parameter. The CS-FEP approach helps to overcome the “endpoint catastrophe” and the “particle collapse” issues suffered by the traditionally used one-step approach. Compared to the three-step approaches, the alchemical process of the combined-structure approach is faster and enable maintaining the neutrality of the system during the whole calculation. The CS-FEP was extensively tested, showing remarkable convergence and satisfactory accuracy. It is also possible to further reduce the computation cost by decreasing the number of λ windows and/or the simulation time of each window. For most of the tested systems, the CS-FEP simulation can give reasonable results even if only one λ window is used for each alchemical path.

Further, the CS-FEP guided structural optimizations were further applied to the discovery of novel potent PDE1 inhibitors. Starting from hit **9** (IC_{50} : 16.8 μ mol/L) discovered by virtual

Table 5 Computationally predicted and experimentally determined mutation-induced changes in binding free energy ($\Delta\Delta G = \Delta\Delta G_{\text{mutant}} - \Delta\Delta G_{\text{wild}}$) of compound **10a** with PDE1. The experimental binding free energy was derived from the obtained IC_{50} by using equation $\Delta G = -RT/nIC_{50}$.

Mutant	$\Delta\Delta G_{\text{exp}}$ (kcal/mol)	$\Delta\Delta G_{\text{cal}}$ (kcal/mol)
Y228A	0.58	0.75
Q427A	1.22	1.17
F398A	1.58	1.44

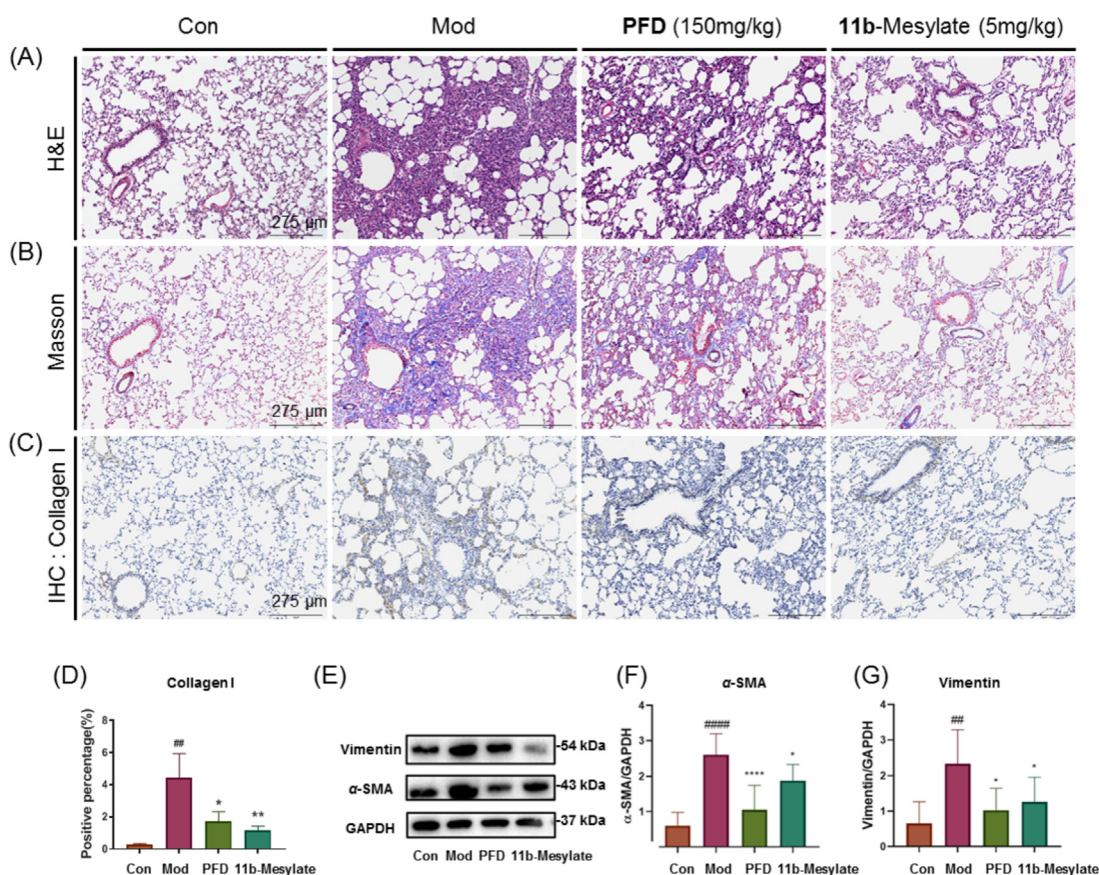


Figure 6 (A) Representative hematoxylin-eosin (H&E) staining sections (magnification $\times 100$) from the lung tissues of the control, BLM-treated, PFD-treated, and **11b**-Mesylate treated groups. The unstained area is alveolar carcinoma, blue is the nucleus, and red is the cytoplasm, respectively. (B) Representative images of Masson's trichrome staining (magnification $\times 100$) from the lung tissues of the control, BLM-treated, PFD-treated, and **11b**-Mesylate treated groups. The unstained area is alveolar carcinoma, dark blue is collagen, and the red stained area is the cell structure. (C) and (D) Representative immunohistochemical staining of Collagen I in lung tissue (magnification $\times 100$) and quantitative analysis of collagen area ($n > 8$). P values are indicated by the non-parametric one-way ANOVA. Compared with the control group, ### $P < 0.01$; Compared with the model group, * $P < 0.05$, ** $P < 0.01$. (E)–(G) Expressions of Vimentin and α -SMA by the Western blot method. p values were indicated by the non-parametric one-way ANOVA. Compared with the control group, ### $P < 0.01$, #### $P < 0.0001$; Compared with the model group, * $P < 0.05$, **** $P < 0.0001$. Data are presented as mean \pm SEM.

screening, the CS-FEP guided hit-to-lead optimization increased the binding affinity by >2000 fold (IC_{50} of compounds **11c** and **11b**: 3.8 and 7.0 nmol/L, respectively). Compound **11b** in its mesylate salt formulation (**11b**-Mesylate) with potent inhibitory activity and considerable metabolic stability further showed remarkable anti-fibrotic effects in BLM-induced IPF rat model.

In short, the theoretical development of the CS-FEP approach and its successful application in the practical drug discovery effort with a new, promising target (PDE1) provides not only a basis for rational discovery of novel anti-fibrotic agents, but also a new and more reliable computational tool for the FEP-based drug design and discovery in general.

4. Experimental

The binding free energy difference of a specific molecular pair A-B can be calculated based on the thermodynamic cycle shown in Fig. 2A. A key point in the CS-FEP approach is the construction of a new alchemical transformation pathway using a combined structure. The combined structure consists of three parts. One part includes the common atoms in both the reference and target

structures, the second part includes the unique atoms in the reference structure, and the third part has the unique atoms in the target structure. Common atoms may have different atom types and vdW parameters, which need to be appropriately accounted for in the FEP calculations. In such cases, the vdW parameters of the common atoms in the combined structure are determined based on the average value of the vdW parameters of the corresponding atoms in both states A and B. The CS-FEP approach simultaneously modifies both the vdW parameters and charges of these atoms during the alchemical simulations. Furthermore, CS-FEP approach addresses the challenge of maintaining system neutrality during the alchemical transformation by implementing a charge redistribution scheme. See details for the treatment of the vdW term and the Coulomb term of the combined structure in Section S1. The binding mode between the receptor and the reference ligand should be known, which can be either a crystal structure or a structure predicted by molecular docking. Briefly, for each alchemical transformation in the thermodynamic cycle, 13 λ windows (0.00, 0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 0.95, and 1.00) were applied to the free energy calculation, resulting in a total of 52 windows for each RBFE

simulation. The detailed method for the CS-FEP molecular dynamic (MD) simulation is given in [Section S1](#).

The free energy differences between the adjacent λ windows were calculated by the BAR (Bennett Acceptance Ratio) estimator⁴⁰. For the ΔU data generated by the MD simulation for each window, the last 1 ns of ΔU was used for the free energy estimation. The total free energy change of each transformation was obtained by accumulating the free energy differences calculated in all the λ windows. Besides, we carried out three kinds of time-convergence analysis, namely time-forward estimate, time-reverse estimate, and time-moving estimate, to check the of the free energy calculation. See the details of time-convergence analysis in [Section S3](#).

The details of preparation and bioassays of the PDE1 inhibitors can be seen in the [Supporting Information, Sections S13, S5](#). ¹H NMR, ¹³C NMR, and high-resolution mass spectra (HRMS); and purity of the tested compounds can be seen in [Supporting Information Sections S14, S15 and S16](#), respectively. One of PDE1 inhibitors (**11b**-Mesylate) was used to study its pharmacological effect based on the bleomycin-induced IPF rat model (see details in [Supporting Information Section S6](#)). All animal care and experimental protocols were in accordance with the “Guide for the Care and Use of Laboratory Animals” (National Institutes of Health Publication, revised 1996, No. 86-23, Bethesda, MD) and approved by the Institutional Ethical Committee for Animal Research of Sun Yat-sen University (SYSU-IACUC-2021-000247).

Acknowledgments

This work was supported by the National Key R&D Program of China (2023YFF1205102), National Natural Science Foundation of China (82273856, 22077143, 21977127, 22377023), the Research Project (31511010402, China), Fundamental Research Funds for Hainan University (KYQD(ZR)-21031, KYQD(ZR)-21108 and XTCX2022JKA01, China), Science Foundation of Hainan Province (KJRC2023B10, China).

Author contributions

Zhe Li: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology. Mei-Yan Jiang: Writing – original draft, Methodology, Investigation. Runduo Liu: Validation, Software, Methodology. Quan Wang: Writing – original draft, Methodology, Data curation. Qian Zhou: Writing – original draft, Validation, Supervision, Methodology, Investigation. Yi-You Huang: Writing – original draft, Validation, Resources, Methodology, Investigation. Yinuo Wu: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology, Investigation. Chang-Guo Zhan: Writing – review & editing, Writing – original draft, Validation. Hai-Bin Luo: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Funding acquisition.

Conflicts of interest

The authors declare no competing financial interest.

Appendix A. Supporting information

Supporting information to this article can be found online at <https://doi.org/10.1016/j.apsb.2024.06.021>.

References

- Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* 2004;**3**:935–49.
- Kuhn B, Gerber P, Schulz-Gasch T, Stahl M. Validation and use of the MM-PBSA approach for drug discovery. *J Med Chem* 2005;**48**:4040–8.
- Wang E, Sun H, Wang J, Wang Z, Liu H, Zhang JZH, et al. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. *Chem Rev* 2019;**119**:9478–508.
- Åqvist J, Medina C, Samuelsson JE. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng Des Sel* 1994;**7**:385–91.
- Hansson T, Marelus J, Åqvist J. Ligand binding affinity prediction by linear interaction energy methods. *J Comput Aided Mol Des* 1998;**12**:27–35.
- Zwanzig RW. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J Comput Phys* 1954;**22**:1420–6.
- Zwanzig RW. High-temperature equation of state by a perturbation method. II. Polar gases. *J Comput Phys* 1955;**23**:1915–22.
- Aldeghi M, Heifetz A, Bodkin MJ, Knapp S, Biggin PC. Accurate calculation of the absolute free energy of binding for drug molecules. *Chem Sci* 2016;**7**:207–18.
- Aldeghi M, Heifetz A, Bodkin MJ, Knapp S, Biggin PC. Predictions of ligand selectivity from absolute binding free energy calculations. *J Am Chem Soc* 2017;**139**:946–57.
- Fu H, Chen H, Cai W, Shao X, Chipot C. BFEE2: automated, streamlined, and accurate absolute binding free-energy calculations. *J Chem Inf Model* 2021;**61**:2116–23.
- Fu H, Chen H, Blazhynska M, Goulard Coderc de Lacam E, Szczepaniak F, Pavlova A, et al. Accurate determination of protein: ligand standard binding free energies from molecular dynamics simulations. *Nat Protoc* 2022;**17**:1114–41.
- Li Z, Li X, Huang YY, Wu Y, Liu R, Zhou L, et al. Identify potent SARS-CoV-2 main protease inhibitors via accelerated free energy perturbation-based virtual screening of existing drugs. *Proc Natl Acad Sci U S A* 2020;**117**:27381–7.
- Li Z, Huang Y, Wu Y, Chen J, Wu D, Zhan CG, et al. Absolute binding free energy calculation and design of a subnanomolar inhibitor of phosphodiesterase-10. *J Med Chem* 2019;**62**:2099–111.
- Li Z, Wu C, Li Y, Liu R, Lu K, Wang R, et al. Free energy perturbation-based large-scale virtual screening for effective drug discovery against COVID-19. *Int J High Perform Comput Appl* 2022;**37**:45–57.
- Wu D, Zheng X, Liu R, Li Z, Jiang Z, Zhou Q, et al. Free energy perturbation (FEP)-guided scaffold hopping. *Acta Pharm Sin B* 2022;**12**:1351–62.
- Liu X, Li Z, Liu S, Sun J, Chen Z, Jiang M, et al. Potential therapeutic effects of dipyridamole in the severely ill patients with COVID-19. *Acta Pharm Sin B* 2020;**10**:1205–15.
- Ma C, Wang J. Dipyridamole, chloroquine, montelukast sodium, candesartan, oxytetracycline, and atazanavir are not SARS-CoV-2 main protease inhibitors. *Proc Natl Acad Sci U S A* 2021;**118**:e2024420118.
- Li Z, Lin Y, Huang YY, Liu R, Zhan CG, Wang X, et al. Reply to Ma and Wang: reliability of various *in vitro* activity assays on SARS-CoV-2 main protease inhibitors. *Proc Natl Acad Sci U S A* 2021;**118**:e2024937118.
- Behnam MAM, Klein CD. Inhibitor potency and assay conditions: a case study on SARS-CoV-2 main protease. *Proc Natl Acad Sci U S A* 2021;**118**:e2106095118.
- Li Z, Liu R, Zhan CG, Wang X, Luo HB. Reply to behnam and klein: potential role of the his-tag in C-terminal his-tagged SARS-CoV-2 main protease. *Proc Natl Acad Sci U S A* 2021;**118**:e2108209118.
- Fighting COVID-19 with HPC. *Nat Comput Sci* 2021;**1**:769–70.
- Wang L, Wu Y, Deng Y, Kim B, Pierce L, Krilov G, et al. Accurate and reliable prediction of relative ligand binding potency in

- prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J Am Chem Soc* 2015;**137**:2695–703.
23. Song LF, Lee TS, Zhu C, York DM, Merz Jr KM. Using AMBER18 for relative free energy calculations. *J Chem Inf Model* 2019;**59**:3128–35.
24. Wong CF, McCammon JA. Dynamics and design of enzymes and inhibitors. *J Am Chem Soc* 1986;**108**:3830–2.
25. Lee TS, Allen BK, Giese TJ, Guo Z, Li P, Lin C, et al. Alchemical binding free energy calculations in AMBER20: advances and best practices for drug discovery. *J Chem Inf Model* 2020;**60**:5595–623.
26. Zacharias M, Straatsma TP, McCammon JA. Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J Chem Phys* 1994;**100**:9025–31.
27. Beutler TC, Mark AE, van Schaik RC, Gerber PR, van Gunsteren WF. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem Phys Lett* 1994;**222**:529–39.
28. Pal RK, Gallicchio E. Perturbation potentials to overcome order/disorder transitions in alchemical binding free energy calculations. *J Chem Phys* 2019;**151**:124116.
29. Jiang W, Chipot C, Roux B. Computing relative binding affinity of ligands to receptor: an effective hybrid single-dual-topology free-energy perturbation approach in NAMD. *J Chem Inf Model* 2019;**59**:3794–802.
30. Rocklin GJ, Mobley DL, Dill KA, Hunenberger PH. Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: an accurate correction scheme for electrostatic finite-size effects. *J Chem Phys* 2013;**139**:184103.
31. Oshima H, Sugita Y. Modified Hamiltonian in FEP calculations for reducing the computational cost of electrostatic interactions. *J Chem Inf Model* 2022;**62**:2846–56.
32. Lee TS, Tsai HC, Ganguly A, York DM. ACES: optimized alchemically enhanced sampling. *J Chem Theor Comput* 2023;**19**:472–9.
33. Lee TS, Lin Z, Allen BK, Lin C, Radak BK, Tao Y, et al. Improved alchemical free energy calculations with optimized smoothstep soft-core potentials. *J Chem Theor Comput* 2020;**16**:5512–25.
34. Tsai HC, Lee TS, Ganguly A, Giese TJ, Ebert MC, Labute P, et al. AMBER free energy tools: a new framework for the design of optimized alchemical transformation pathways. *J Chem Theor Comput* 2023;**19**:640–58.
35. Wu Y, Tian YJ, Le ML, Zhang SR, Zhang C, Huang MX, et al. Discovery of novel selective and orally bioavailable phosphodiesterase-1 inhibitors for the efficient treatment of idiopathic pulmonary fibrosis. *J Med Chem* 2020;**63**:7867–79.
36. Huang MX, Tian YJ, Han C, Liu RD, Xie X, Yuan Y, et al. Structural modifications of nimodipine lead to novel PDE1 inhibitors with anti-pulmonary fibrosis effects. *J Med Chem* 2022;**65**:8444–55.
37. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph* 1996;**14**:33–8.
38. Lu T, Chen F. Multiwfn: a multifunctional wavefunction analyzer. *J Comput Chem* 2011;**33**:580–92.
39. Loeffler HH, Bosio S, Duarte Ramos Matos G, Suh D, Roux B, Mobley DL, et al. Reproducibility of free energy calculations across different molecular simulation software packages. *J Chem Theor Comput* 2018;**14**:5567–82.
40. Bennett CH. Efficient estimation of free energy differences from Monte Carlo data. *J Comput Phys* 1976;**22**:245–68.
41. Li Y, Liu R, Liu J, Luo H, Wu C, Li Z. An open source graph-based weighted cycle closure method for relative binding free energy calculations. *J Chem Inf Model* 2023;**63**:561–70.
42. Hahn DF, Bayly CI, Boby ML, Bruce Macdonald HE, Chodera JD, Gapsys V, et al. Best practices for constructing, preparing, and evaluating protein–ligand binding affinity benchmarks [Article v1.0]. *Living J Comp Mol Sci* 2022;**4**:1497.
43. Drake TM, Docherty AB, Harrison EM, Quint JK, Adamali H, Agnew S, et al. Outcome of hospitalization for COVID-19 in patients with interstitial lung disease. an international multicenter study. *Am J Respir Crit Care Med* 2020;**202**:1656–65.
44. George PM, Wells AU, Jenkins RG. Pulmonary fibrosis and COVID-19: the potential role for antifibrotic therapy. *Lancet Respir Med* 2020;**8**:807–15.
45. Lechowicz K, Drozdal S, Machaj F, Rosik J, Szostak B, Zegan-Baranska M, et al. COVID-19: the potential treatment of pulmonary fibrosis associated with SARS-CoV-2 infection. *J Clin Med* 2020;**9**:1917.
46. Wendisch D, Dietrich O, Mari T, von Stillfried S, Ibarra IL, Mittermaier M, et al. SARS-CoV-2 infection triggers profibrotic macrophage responses and lung fibrosis. *Cell* 2021;**184**:6243–6261.e27.
47. Hirawat R, Jain N, Aslam Saifi M, Rachamalla M, Godugu C. Lung fibrosis: post-COVID-19 complications and evidences. *Int Immunopharmacol* 2023;**116**:109418.
48. Li Z, Cai YH, Cheng YK, Lu X, Shao YX, Li X, et al. Identification of novel phosphodiesterase-4D inhibitors prescreened by molecular dynamics-augmented modeling and validated by bioassay. *J Chem Inf Model* 2013;**53**:972–81.
49. Huang YY, Yu YF, Zhang C, Chen Y, Zhou Q, Li Z, et al. Validation of phosphodiesterase-10 as a novel target for pulmonary arterial hypertension via highly selective and subnanomolar Inhibitors. *J Med Chem* 2019;**62**:3707–21.
50. Huang YY, Deng J, Tian YJ, Liang J, Xie X, Huang Y, et al. Mangostanin derivatives as novel and orally active phosphodiesterase 4 inhibitors for the treatment of idiopathic pulmonary fibrosis with improved safety. *J Med Chem* 2021;**64**:13736–51.
51. Song Z, Huang YY, Hou KQ, Liu L, Zhou F, Huang Y, et al. Discovery and structural optimization of toddacoumalone derivatives as novel PDE4 inhibitors for the topical treatment of psoriasis. *J Med Chem* 2022;**65**:4238–54.
52. Wu XN, Zhou Q, Huang YD, Xie X, Li Z, Wu Y, et al. Structure-based discovery of orally efficient inhibitors via unique interactions with H-pocket of PDE8 for the treatment of vascular dementia. *Acta Pharm Sin B* 2022;**12**:3103–12.
53. Zhou Q, Le M, Yang Y, Wang W, Huang Y, Wang Q, et al. Discovery of novel phosphodiesterase-1 inhibitors for curing vascular dementia: suppression of neuroinflammation by blocking NF- κ B transcription regulation and activating cAMP/CREB axis. *Acta Pharm Sin B* 2023;**13**:1180–91.