ORIGINAL ARTICLE

AncPhore: A versatile tool for anchor pharmacophore steered drug discovery with applications in discovery of new inhibitors targeting metallo-β-lactamases and indoleamine/ tryptophan 2,3-dioxygenases

Qingqing Dai[a,†], Yuhang Yan[a,†], Xiangli Ning[a], Gen Li[a], Junlin Yu[a], Ji Deng[a], Lingling Yang[b], Guo-Bo Li[a,*]

[a]Key Laboratory of Drug-Targeting and Drug Delivery System of the Education Ministry and Sichuan Province, Department of Medicinal Chemistry, West China School of Pharmacy, Sichuan University, Chengdu 610041, China
[b]College of Food and Bioengineering, Xihua University, Chengdu 610039, China

Abstract We herein describe AncPhore, a versatile tool for drug discovery, which is characterized by pharmacophore feature analysis and anchor pharmacophore (*i.e.*, most important pharmacophore features) steered molecular fitting and virtual screening. Comparative analyses of numerous protein—ligand complexes using AncPhore revealed that anchor pharmacophore features are biologically important, commonly associated with protein conservative characteristics, and have significant contributions to

Indoleamine 2,3-
dioxygenase;
Tryptophan 2,3-
dioxygenase

the binding affinity. Performance evaluation of AncPhore showed that it had substantially improved prediction ability on different types of target proteins including metalloenzymes by considering the specific contributions and diversity of anchor pharmacophore features. To demonstrate the practicability of AncPhore, we screened commercially available chemical compounds and discovered a set of structurally diverse inhibitors for clinically relevant metallo-β-lactamases (MBLs); of them, **4** and **6** manifested potent inhibitory activity to VIM-2, NDM-1 and IMP-1 MBLs. Crystallographic analyses of VIM-2:**4** complex revealed the precise inhibition mode of **4** with VIM-2, highly consistent with the defined anchor pharmacophore features. Besides, we also identified new hit compounds by using AncPhore for indoleamine/tryptophan 2,3-dioxygenases (IDO/TDO), another class of clinically relevant metalloenzymes. This work reveals anchor pharmacophore as a valuable concept for target-centered drug discovery and illustrates the potential of AncPhore to efficiently identify new inhibitors for different types of protein targets.

1. Introduction

Efficient discovery of new hit/lead compounds for a specific target protein still is an important task in innovative drug research & development (R&D)[1]. As one of the classic computer-assisted drug design methods, pharmacophore approaches have been widely used within the field of drug discovery[2], which are mainly based on the concept that specific molecular interactions are observed in molecular recognition of a ligand by a biological target. Typically, a pharmacophore model is defined as an ensemble of chemical interactions (*e.g.*, hydrogen bonds, charges, and hydrophobic contacts) and their spatial arrangement, which enables the rationalization of interaction patterns for a target protein or a set of chemically diverse active ligands, and the subsequent application for virtual screening, *de novo* design, and lead optimization[2]. Many pharmacophore modelling programs have been established[3], of which some are free for academic users such as Pharao[4] and Pharmer[5], and have been successfully used in a variety of drug discovery projects. Yet despite these successes, pharmacophore approaches have not achieved their expected full potential, particularly in facing the actual demand with regard to reduce current expensive overall drug R&D costs[2].

The effectiveness and efficiency of pharmacophore approaches depend largely on the quality of established pharmacophore model and the understanding of the context and significance of each pharmacophore feature. Usually, a pharmacophore model consists of multiple pharmacophore features to represent the specificity of the target protein to recognize structurally diverse ligands. In practical applications, we noted that a very small proportion of the pharmacophore features have most critical contributions in protein−ligand recognition; we referred such features as 'anchor pharmacophore features'. For example, most clinically useful kinase inhibitors are found to form 2−3 critical hydrogen bonds (*i.e.*, anchor pharmacophore features) with the active-site hinge region, which is a widely recognized characteristic for kinase inhibitor discovery and development[6]. We recently observed that several clinically relevant metallo-β-lactamases, albeit with large sequence differences, contain a specific conservative active-site region involving the anchor pharmacophore features that specifically recognize the β-lactam antibacterials as well as the potent inhibitors[7,8]; notably, the compounds have less potent metallo-β-lactamase inhibitory activity if they do not fit well with the anchor pharmacophore features[7,9]. Although the importance of anchor

pharmacophore features has been gradually recognized, comprehensive analysis and proper application of such features has been lacking until now.

Herein, we report a versatile tool for pharmacophore-based drug discovery, termed AncPhore. It involves fine-tuned methods to analyze pharmacophore features for protein−ligand complex, ligand or apo-protein, particularly including the relatively complicated metal-coordination feature that was not well-described previously. By large-scale cross-target analyses of pharmacophore features, we found that anchor pharmacophore features are commonly related with the conservative characteristics within the family of target proteins, have important biological significance (*e.g.*, involving in catalytic nature or binding of natural ligands), and have great contributions to the ligand binding affinity, which offer relatively comprehensive, conceptual understanding of anchor pharmacophore features. More distinctly, a new algorithm and a new scoring function were introduced in AncPhore to achieve anchor pharmacophore steered molecular fitting and virtual screening. Performance evaluations on different types of protein targets using the DUD-E dataset[10] revealed that AncPhore had remarkably improved prediction ability when considering the distinctive contributions and diversity of anchor pharmacophore features. Applications of AncPhore led to the discovery of new structurally diverse inhibitors for two classes of clinically relevant metalloenzymes: metallo-β-lactamases (MBLs) and indoleamine 2,3-dioxygenase 1 (IDO1)/tryptophan 2,3-dioxygenase (TDO), which further demonstrated the power of AncPhore particularly in hit/lead discovery. The AncPhore program is freely available on the website https://ancphore.ddtmlab.org.

2. Results and discussion

2.1. Anchor pharmacophore revealed by cross-target analysis

AncPhore involves ten types of pharmacophore features, including hydrogen bond donor (HD), hydrogen bond acceptor (HA), positively charged center (PO), negatively charged center (NE), metal coordination (MB), halogen bonding (XB), aromatic ring (AR), cation−π interaction (CR), hydrophobic (HY), and covalent bonding (CV) features (as graphically shown in Supporting Information Fig. S1). Uniquely, MB, XB, and CR features that are

A

Relation of binding affinity and number of pharmacophore features

$F(15, 2198.70) = 50.15$, $P \leq 0.001$, $\widehat{\eta_p^2} = 0.05$, $CI_{95\%}$ [0.04, 0.05], $n_{obs} = 15{,}529$

B

C

Comparison of binding affinity with/without anchor pharmacophore features
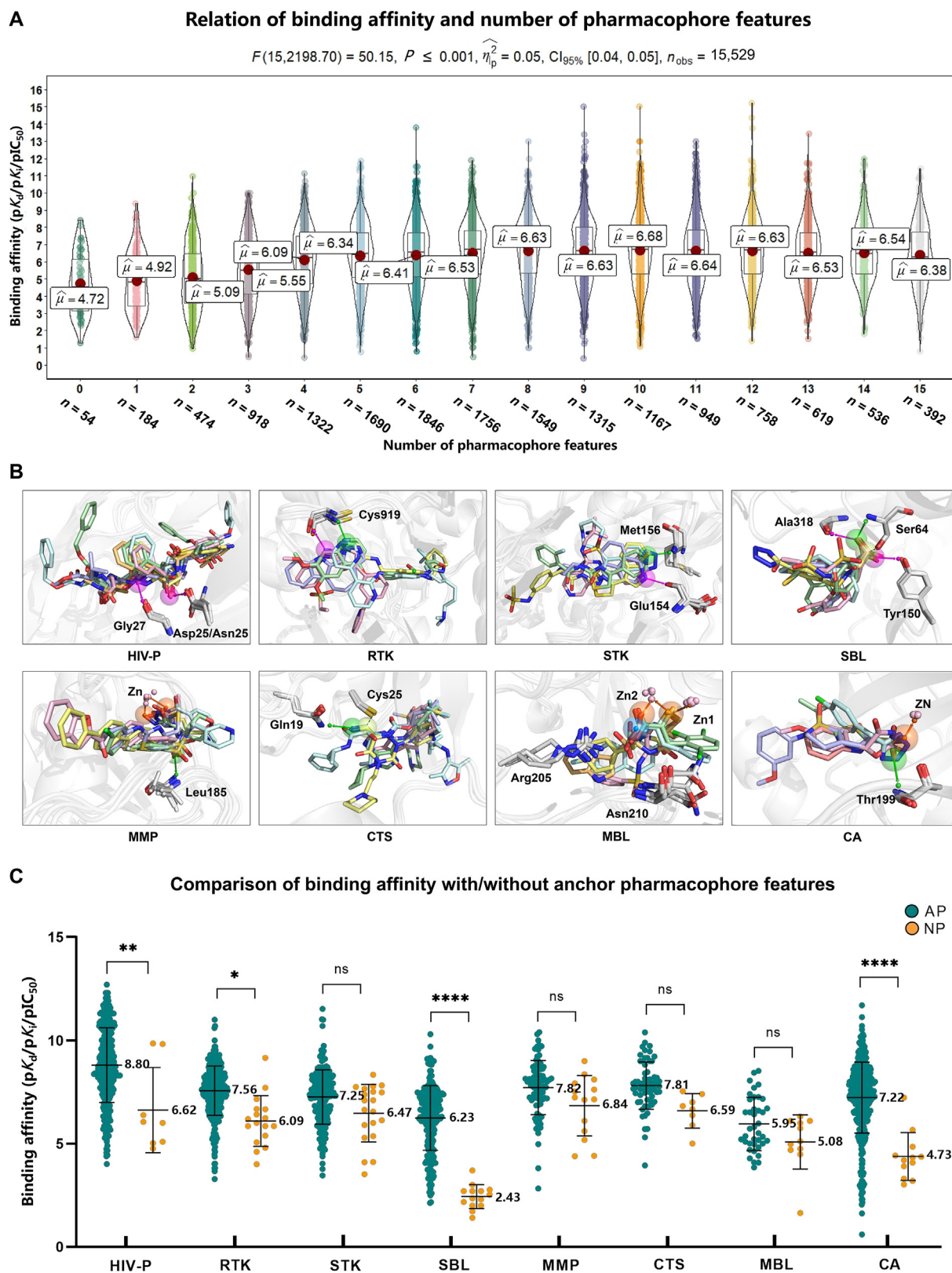
Figure 1 Anchor pharmacophore features are conservative in target protein family and great contributions to the binding affinity. (A) Analysis of number of pharmacophore features *versus* binding affinity of >15,000 complex structures revealed a certain degree of relevance between them. (B) Anchor pharmacophore features are commonly conservative within the family of target proteins, as observed for HIV-P, RTK, STK, SBL, MMP, CTS, MBL, and CA family proteins. (C) Comparison of the binding affinity of the protein−ligand complexes with/without anchor pharmacophore features (AP/NP) indicated the features are of significant importance to achieve the high binding affinity.

not considered in most pharmacophore modelling tools were described in detail (Experimental Section). In particular, the MB feature is defined as to resemble hydrogen bonding interactions, *i.e.*, ligand electron donor corresponding to metal ion acceptor with a proper direction (Fig. S1), which is different from that defined in most pharmacophore modelling programs[3]. The optimal distance and angle for MB in different situations are set according to our analyses of metalloenzyme−ligand complexes[11]. The current version of AncPhore can be used to analyze pharmacophore features for a given protein−ligand complex, a ligand or an apo-protein, as described in the Experimental Section.

By using AncPhore, we carried out pharmacophore feature analyses for more than 17,000 protein−ligand complex structures covering >3600 target entries from PDBbind v2019[12−14]. As shown in Fig. 1A, most protein−ligand complexes had a range from 3 to 15 pharmacophore features, mainly concentrated in 4−10 features. Globally, the binding affinity increased as the number of pharmacophore features increases only when less than 8 pharmacophore features (Fig. 1A). There was no obvious relevance of the number of individual hydrogen bonding, aromatic ring, and metal coordination features with the binding affinity data, with the exception of ionic interaction features (Supporting Information Fig. S2). By comparison, for single target, we can observe a relatively obvious correlation between the binding affinity and the number of pharmacophore features, as analyzed for the targets with >100 complex structures (Supporting Information Fig. S3), suggesting that the pharmacophore features recognized by AncPhore could at least partly reflect the binding affinity content.

We then attempted to analyze anchor pharmacophore features (*i.e.*, most important pharmacophore features) for the target proteins which have a large number of protein−ligand complex structures, including carbonic anhydrase 2 (CA2), human immunodeficiency virus type 1 protease (HIV1-P), and cyclin-dependent kinase 2 (CDK2). Through structure superimposition and pharmacophore feature analysis, for CA2, we observed two most frequent pharmacophore features: a metal coordination feature (MB) and a hydrogen-bonding acceptor (HA1), which exist in 89.28% and 84.29% of 401 analyzed complex structures, respectively (Supporting Information Fig. S4A), while other features are with the frequency low than 30%; consistent with previous studies, MB and HA1 are quite important for inhibitor binding and hence can be recognized as anchor pharmacophore features[15]. Similarly, for HIV1-P and CDK2, we can also observe 2−3 anchor pharmacophore features that are common and important for the binding of most inhibitors (Fig. S4B and S4C).

We next set up a protocol to carry out cross-target analysis of pharmacophore features, with the aim to understand anchor pharmacophore features beyond a single target. In this protocol, all complex structures were first superimposed each other by using a non-sequential structure alignment algorithm[16], followed by grouping structures corresponding to their structural similarity and determining the anchor pharmacophore features. The analyses resulted in 77 target protein groups which contain at least three structurally similar proteins; the proteins in each group mostly belong to the same protein family. Comparison of the complex structures in the top-ranked protein groups revealed that anchor pharmacophore features are commonly conservative within the family of target proteins. As for the HIV-P group, two hydrogen-bonding donors associating with Gly27 and Asp25/Asn25 were recognized as anchor pharmacophore features for the binding of most inhibitors (Fig. 1B), notably, which are directly involved in

substrate hydrolysis reactions (Supporting Information Fig. S5A)[17]. The target proteins in receptor tyrosine kinase (RTK) and serine/threonine kinase (STK) group were found to have common anchor pharmacophore features, *i.e.*, 2−3 hydrogen bonds associating with the active-site hinge region (Fig. 1B), which are critical for binding of the co-substrate ATP adenine core (Fig. S5B and S5C)[18,19]. Similar scenarios were observed for other analyzed target groups, such as serine-β-lactamase (SBL), matrix metalloproteinase (MMP), cathepsins (CTS), metallo-β-lactamase (MBL), and carbonic anhydrase (CA) groups, in which anchor pharmacophore features are associated with the protein conservative characteristics and involved in catalytic nature or binding of natural substrates (Fig. 1B and Fig. S5D−S5H).

We next analyzed how important anchor pharmacophore features are to the binding affinity. All protein−ligand complex structures in each above-analyzed target group were divided into two categories: AP and NP, corresponding to the complex structures with and without anchor pharmacophore features, respectively. We observed substantial difference in binding affinity (1 to 3 orders of magnitude) between AP and NP for all analyzed groups (Fig. 1C). As for HIV-P group, the average of the binding affinity ($pK_i/pK_d/pIC_{50}$) of the protein−ligand complexes in AP is 8.80, while the average in NP is 6.62 (Fig. 1C); for RTK group, the average of the binding affinity in AP and NP are 7.56 and 6.09, respectively (Fig. 1C). These results largely indicated that anchor pharmacophore features have great contributions to the binding affinity.

The overall results provided relatively comprehensive, conceptual understanding of anchor pharmacophore features, namely, that such features are biologically important, commonly associated with protein conservative characteristics, and have significant contributions to the binding affinity. Anchor pharmacophore features are not only able to reflect the nature of biological target proteins evolved to conservatively recognize their natural substrates but also provide key pointcuts for the discovery of physiologically active compounds and drug molecules. In general, anchor pharmacophore features could be identified by pharmacophore feature analysis of sufficient structural information or by investigation of active-site characteristics, *e.g.*, catalytic nature or binding features of natural substrates. The latter manner is previously undervalued and particularly important for the target proteins whose complex structures are rarely available. Although anchor pharmacophore features are focused on protein−ligand interactions in this study, such features should be not confined to protein−ligand recognition. For example, in protein−protein recognition, the 'hot-spot' residues are usually observed, which provide key features for specific protein−protein recognition as well as identification of drug-like small molecules to alter the protein−protein recognition[20]. Analysis of anchor pharmacophore features in protein−protein recognition is worth expecting. Given the importance of anchor pharmacophore features, appropriate treatment of such features in virtual screening, *de novo* molecular design, and structural optimization is most likely to improve the efficiency of drug discovery.

2.2. Anchor pharmacophore steered virtual screening

We next examined the potential of AncPhore in virtual screening by consideration of the specific contributions of anchor pharmacophore features. In AncPhore, a sophisticated scoring function called APScore was introduced to evaluate the fitness of queried molecules with a given pharmacophore model; APScore
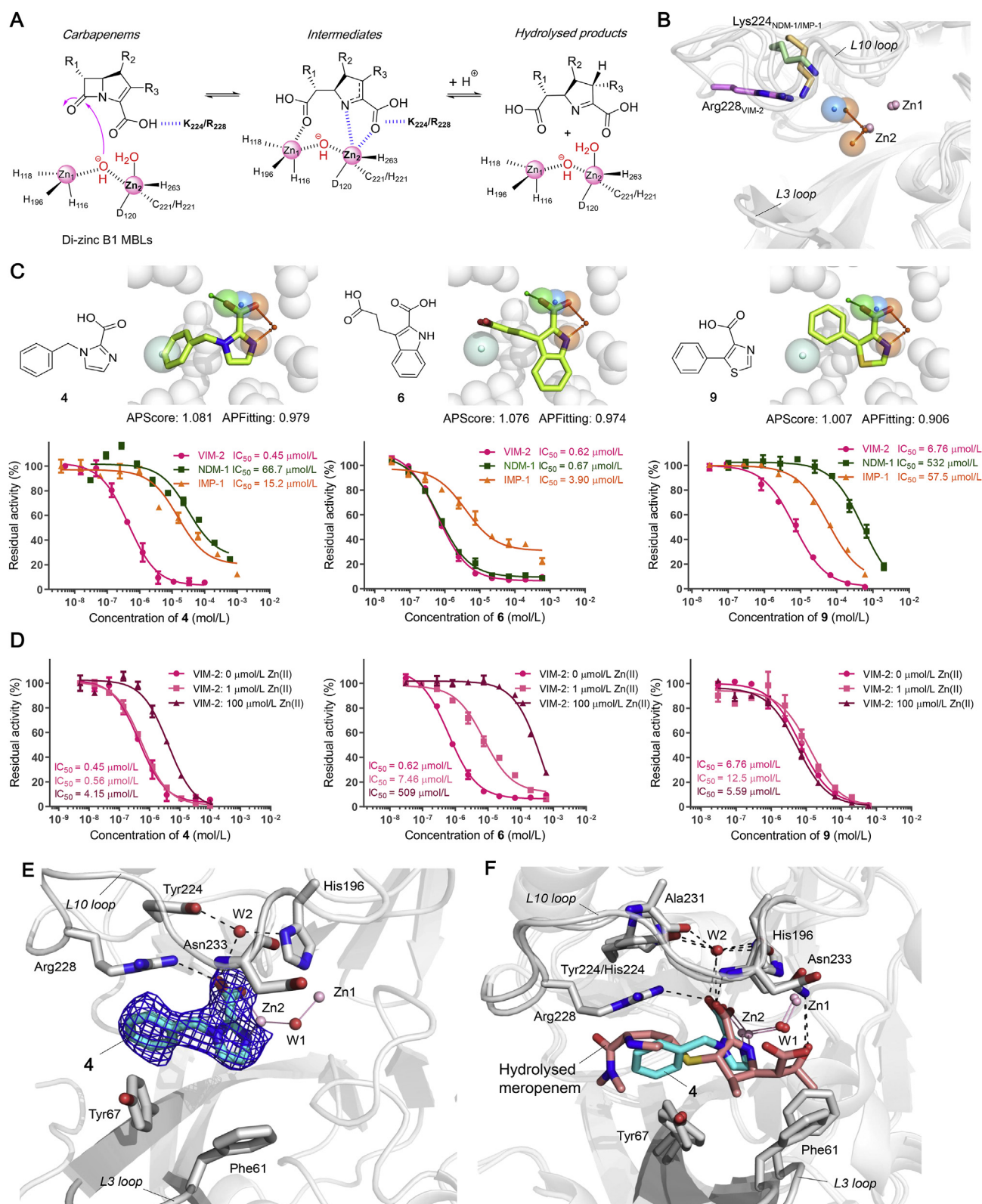
Figure 2 The activity and mode-of-action of new MBL inhibitors identified by AncPhore. (A) The catalytic mechanism of class B1 di-zinc MBLs. (B) Defined anchor pharmacophore features for B1 MBLs. (C) The inhibitory activity of **4**, **6**, and **9** with VIM-2, NDM-1, and IMP-1 and their fitting modes with the pharmacophore models. (D) The inhibitory activity for **4**, **6**, and **9** with VIM-2 at three different concentrations of Zn(II) (0, 1, and 100 μmol/L) revealed their difference in the potential of zinc ion chelation in solution. (E) Crystallographic analysis revealed the binding mode of **4** with VIM-2; the mF_o-DF_c electron density (OMIT maps) around **4** (blue mesh, contoured to 3σ) and view from a crystal structure of the VIM-2:**4** complex (PDB code 7CHV). (F) Superimposition of VIM-2:**4** and VIM-1:hydrolyzed meropenem (PDB code 5N5I)[21] revealed the importance of anchor pharmacophore features and their intrinsic connection with the substrate binding nature.

characteristically involves the inherent weights for different types of pharmacophore features and especially resizable weights for anchor pharmacophore features (Experimental Section). More uniquely, AncPhore is developed to achieve anchor pharmacophore guided molecular fitting process and final ranking of screened molecules (Experimental Section).

We first evaluated the performance of AncPhore on eight target proteins (including HIV1-P, VEGFR2, ROCK1, AMPC, MMP13, CatK, VIM-2, and CA2) with or without consideration of anchor pharmacophore features for a single pharmacophore model that was generated from an apo-structure or a complex structure. The test dataset for each target was downloaded or generated from DUD-E[10]. Almost for all the tested targets, the area under curve (AUC) values for a single model (either from an apo-structure or from a complex structure) with consideration of anchor pharmacophore features are higher than that without setting the anchor pharmacophore features (Supporting Information Fig. S6); meanwhile, anchor pharmacophore features were observed to play important roles in enhancing enrichment factors in virtual screening (Supporting Information Table S1). We also observed that AncPhore has better performance with average AUC of 0.80 when using the pharmacophore models from complex structures than that from apo-structures (average AUC = 0.69, Fig. S6), possibly implying that the former can better reflect the targets' features.

Given the flexibility and dynamics of binding sites particularly for metalloenzymes, we then tested the performance of AncPhore by simultaneously considering the specific contributions and diversity of anchor pharmacophore features (i.e., using multiple pharmacophore models). AncPhore manifested markedly improved prediction ability on the tested target proteins with the AUC values ranging from 0.80 to 0.865 (Fig. S6) and the enrichment factors at 1% ranging from 2.27 to 19.15 (Table S1). Notably, for the metalloenzymes MMP13, VIM-2, and CA2, the AUC values by multiple-model manner are 0.865, 0.858, and 0.863, respectively, which are higher than that by single-model manner (Fig. S6). These results suggest that appropriate use of multiple pharmacophore features, even for metalloenzymes which commonly involve complicated anchor metal coordination features (Fig. 1B), will be able to largely improve performance in virtual screening; this constitutes the distinctive advantage of anchor pharmacophore guided strategy in drug discovery.

2.3.   Application cases

We subsequently examined the power of AncPhore in virtual screening for MBLs. MBLs are β-lactamases that use zinc ions to activate a nucleophilic water molecule to hydrolyze almost all β-lactam antibiotics (Fig. 2A), including the 'last-resort' carbapenems, which represent one class of most attractive targets to combat antibacterial resistance[7,8]. The current research focus is to identify new inhibitor chemotypes for broad-spectrum targeting the most clinically relevant MBLs, such as NDM, VIM, and IMP types[7]. Analyses of reported complex structures for various MBL enzymes revealed the anchor pharmacophore features, i.e., a negatively charged center and metal coordination features, corresponding to the positively charged residues (e.g., K224$_{NDMs/IMPs}$ and R228$_{VIMs}$) and the second active-site zinc ion (Zn$_2$), respectively (Fig. 2B), which have been demonstrated as essential features for β-lactam substrate recognition[7]. Due to crystallographically observed multiple coordination modes with the active-site zinc ions, we selected six distinct pharmacophore

models with defined anchor pharmacophore features (Supporting Information Fig. S7) to screen commercially available compounds with molecular weights less than 350 in the VITAS-M and enamine library. From the top-ranked hit compounds, compounds 1−10 (Table 1) were chosen for experimental validation, since their chemical scaffolds have not been reported as MBL inhibitors to date. The fitting modes and APScore values of 1−10 with the pharmacophore models are shown in Fig. 2C and Supporting Information Fig. S8.

As shown in Table 1, compounds 1−10 displayed varying degrees of inhibitory activity against the representative MBL enzymes: VIM-2, NDM-1, and IMP-1. Of them, four compounds manifested potent inhibitory activity to one or all of the tested enzymes, including 3, 4, 6, and 9. Compound 3 displayed IC$_{50}$ values of 38.15, 196.0, and 2.37 μmol/L to VIM-2, NDM-1, and IMP-1, respectively. Compound 4 had more potent inhibition than L-captopril (Table 1) against VIM-2 and NDM-1 with IC$_{50}$ values of 0.45 and 66.7, respectively (Fig. 2C); the ligand efficiency (LE) values of 4 to VIM-2, NDM-1, and IMP-1 are 0.59, 0.39, and 0.45, respectively (Table 1). The LE value calculated using Eq. (1):

$$LE = 1.4 \times pIC_{50}/N \tag{1}$$

where N is the number of heavy atoms (i.e., non-hydrogen atoms). Notably, compound 6 manifested potent broad-spectrum inhibition to all the tested enzymes (Fig. 2C); the IC$_{50}$ values to VIM-2, NDM-1, and IMP-1 are 0.62 μmol/L (LE = 0.51), 0.67 μmol/L (LE = 0.51), and 3.90 μmol/L (LE = 0.45), respectively (Table 1). Compound 9 showed good inhibitory activity to VIM-2 (IC$_{50}$ = 6.76 μmol/L) and moderate activity to NDM-1 (IC$_{50}$ = 532 μmol/L) and IMP-1 (IC$_{50}$ = 57.5 μmol/L, Table 1). We further tested the inhibitory activity for 4, 6, and 9 with VIM-2 at three different concentrations of Zn(II) (0, 1, and 100 μmol/L) to examine the potential of zinc ion chelation in solution. With addition of 100 μmol/L zinc ions, compound 4 showed relatively low inhibitory activity to VIM-2 compared with the low concentrations of zinc ions (Fig. 2D). Unexpectedly, the inhibitory activity of 6 to VIM-2 decreased significantly with increasing concentrations of zinc ions (Fig. 2D), suggesting that 6 has strong ability to chelate with zinc ions in solution and may cause non-specific influence on metal ions containing species. In contract, there is no obvious differences between the inhibitory activity of 9 with or without excess zinc ions (Fig. 2D), indicating that 9 is not a strong zinc ion chelator in solution. We then tested whether 4, 6, and 9 have potential to reverse resistance of meropenem, a representative carbapenem antibiotic, in Escherichia coli with expression of VIM-2 MBL (Experimental Section). Compound 4 showed more potent activity than 9 in the cell-based assays, which could reduce MIC values of meropenem to 0.125 μg/mL (32 fold) and 0.25 μg/mL (16 fold) when treatment of 100 μg/mL and 10 μg/mL 4, respectively (Supporting Information Table S2). By comparison, compound 6, albeit with potent enzymatic activity, had no obvious cellular activity (Table S2), probably due to its poor cell permeability or other factors.

We then carried out co-crystallization experiments for VIM-2 in complex with 4, which has a highest LE value (LE = 0.59) to VIM-2 (Table 1) and good cellular activity (Table S2). The VIM-2:4 complex structure (PDB code 7CHV) was solved to 2.17 Å resolution, in which 4 was well-modelled with the clearly defined electron density in the active site (Supporting Information Table S3 and Fig. 2E). In VIM-2:4 structure, we observed that the

Table 1 The inhibitory activity (IC_{50}) and ligand efficiency (LE) of compounds **1**–**10** with clinically relevant VIM-2, NDM-1, and IMP-1.

Compd.	Chemical structure[a]	IC_{50} (μmol/L)[b]/LE		
		VIM-2	NDM-1	IMP-1
1		>1000/<0.30	>1000/<0.30	>1000/<0.30
2		540.0/0.27	>1000/<0.25	347.8/0.28
3		38.15/0.31	196.0/0.26	2.37/0.39
4		0.45/0.59	66.7/0.39	15.2/0.45
5		168.7/0.31	>1000/<0.25	266.9/0.29
6		0.62/0.51	0.67/0.51	3.90/0.45
7		402.2/0.32	491.2/0.31	49.75/0.40
8		857.9/0.29	>1000/<0.28	398.4/0.32
9		6.76/0.52	532/0.33	57.5/0.42
10		962.3/0.28	>1000/<0.28	525.8/0.31
L-cap[c]		2.47/0.56	196.8/0.37	2.50/0.56

[a]Compounds **1**–**10** were confirmed by high-resolution mass spectrometry (Supporting Information) and used without further purification.
[b]The IC_{50} values ($n \geq 3$) were measured as described in Experimental Section, and the IC_{50} curves were given in Fig. 2 and Fig. S8.
[c]The IC_{50}/pIC_{50} values of L-captopril are from Ref. 22 which were obtained under the same assay conditions.

carboxyl oxygen and N-3 atom of the imidazole ring of **4** are positioned to chelate with the active site Zn2, and its carboxyl group forms electrostatic interactions with Arg228 on the L10 loop, and its phenyl group makes face-to-edge π–π stacking and hydrophobic interactions with Tyr67 on the L3 loop (Fig. 2E). The crystal structure fully demonstrated that **4** binds to form proposed anchor pharmacophore features with Zn2 and Arg228 (Fig. 2E). Superimposing VIM-2:**4** with VIM-1:hydrolyzed meropenem (PDB ID 5N5I)[21] structures revealed that **4** mimics the binding of the β-lactam substrate, ring-opened meropenem, e.g., positioned to chelate with Zn2 via C-2 carboxylate group and exposed N-1

atom, and make hydrogen-bonding interactions with Arg228/Asn233 (Fig. 2F); this further indicated the intrinsic connection of anchor pharmacophore features in substrate and inhibitor binding.

Subsequently, we used AncPhore for virtual screening against human IDO/TDO enzymes, which are heme-containing dioxygenases that catalyze the oxidative cleavage of the 2,3-indole position of tryptophan (Trp) to generate N-formyl kynurenine (Fig. 3A). These enzymes are first and rate-limiting enzymes of the kynurenine pathway, and implicated in neurodegenerative diseases and tumoral immune resistance; in particular, IDO1 and TDO have been demonstrated as important targets for cancer
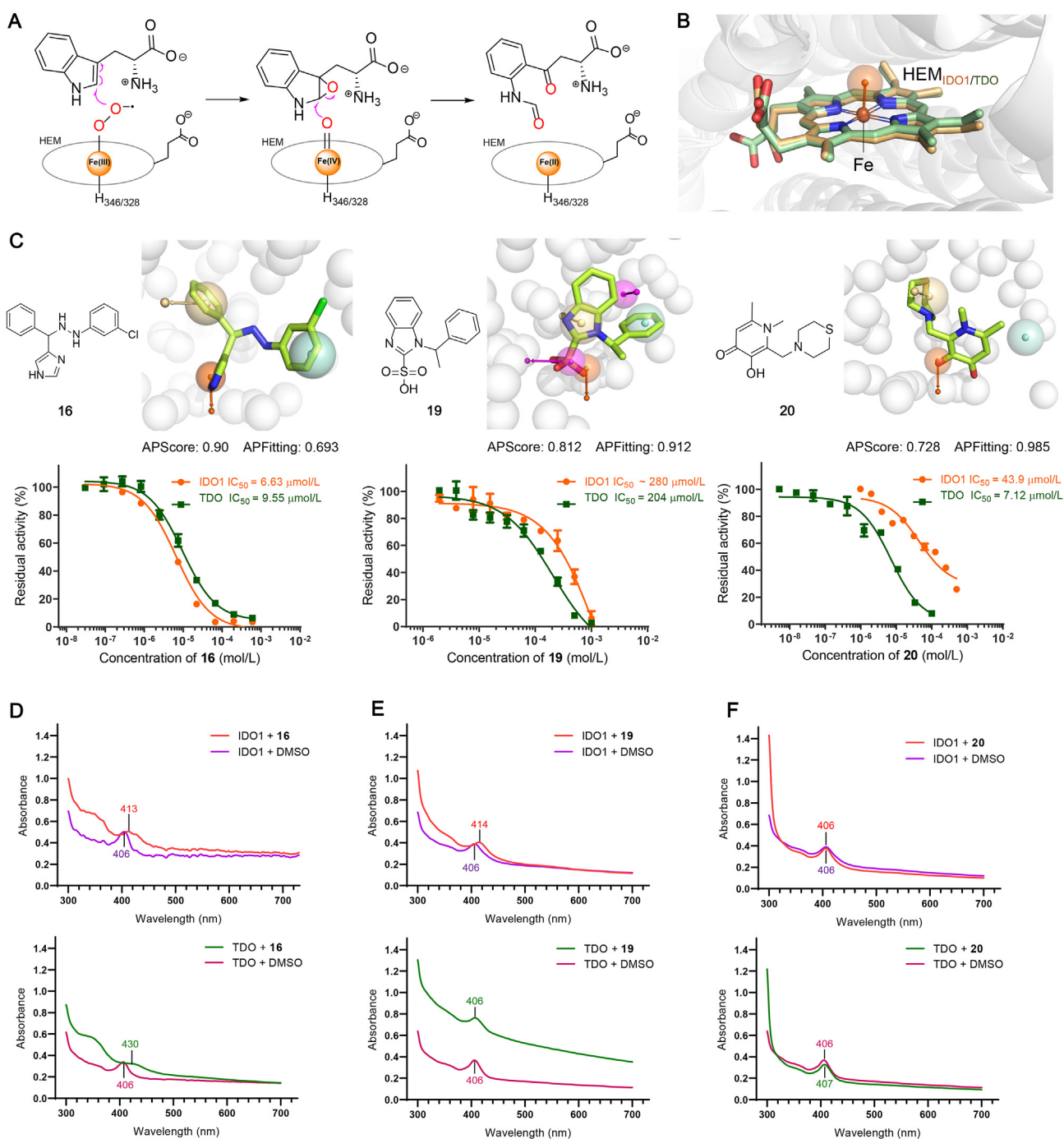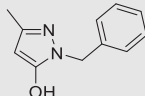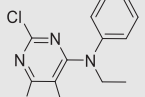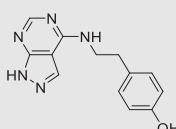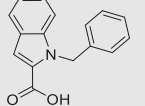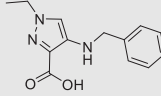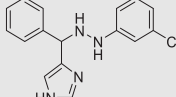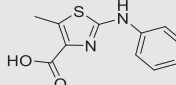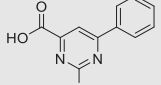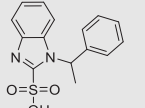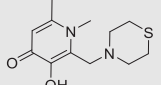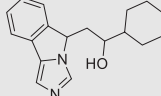
Figure 3 The inhibitory activity and binding features of hit compounds **16**, **19**, and **20** with IDO1/TDO. (A) The catalytic mechanisms of IDO1/TDO mediated oxidative cleavage of the 2,3-indole position of Trp. (B) A metal coordination feature is defined as the anchor pharmacophore feature. (C) The fitting modes of **16**, **19**, and **20** with the pharmacophore models and their IC$_{50}$ curves with IDO1/TDO. (D)–(F) UV–Vis absorption spectra of IDO1/TDO in complex with **16**, **19**, or **20**, revealing that **16** and **19** are likely to bind with heme-iron of IDO1, while **20** does not involve the heme binding for both IDO1 and TDO.

immunotherapy[23−26]. Although a number of IDO1 and TDO inhibitors have been reported and several inhibitors are in preclinical development[26−29], discovery of new inhibitor chemotypes is still desirable at present. By comprehensive analysis of all reported complex structures for IDO1/TDO and the catalytic mechanisms of the tryptophan substrate (Fig. 3A), we defined an anchor metal-coordination pharmacophore feature (Fig. 3B) and selected 4 pharmacophore models for virtual screening (Supporting

Information Fig. S9). By visual inspection and novelty searching, we selected 10 structurally different hit compounds (**11−20**, Table 2) for biological tests from the top-ranked compounds identified by AncPhore, and their fitting modes with the pharmacophore models are shown in Fig. 3C and Fig. S10.

As shown in Table 2, several compounds showed inhibition to IDO1 and/or TDO, of which **16** and **20** are most potent inhibitors for IDO1 and TDO. **16** displayed IC$_{50}$ values of 6.63 μmol/L

Table 2 The IC$_{50}$ and LE values of hit compounds **11−20** with IDO1/TDO.

Compd.	Chemical structure[a]	IC$_{50}$ (μmol/L)[b]/LE[c]	
		*h*IDO1	*h*TDO
11		>500/<0.33	268/0.36
12		>1000/<0.24	237.4/0.27
13		>500/<0.24	201.9/0.27
14		>500/<0.24	>500/<0.24
15		∼300/∼0.27	262.3/0.28
16		6.63/0.35	9.55/0.33
17		261.7/0.31	>500/<0.29
18		>500/<0.29	∼300/∼0.31
19		∼280/0.24	204/0.25
20		43.9/0.36	7.12/0.42
NLG-919 analogue		0.09/0.47	0.15/0.45

[a]Compounds **11−20** were confirmed by high-resolution mass spectrometry (Supporting Information) and used without further purification.
[b]The method for measuring IC$_{50}$ values ($n \geq 3$) is described in Experimental Section; IC$_{50}$ curves are shown in Fig. 3 and Supporting Information Fig. S10.
[c]The LE values are calculated as Eq. (1).

(LE = 0.35) and 9.55 μmol/L (LE = 0.33) to IDO1 and TDO, respectively; **20** had activity of 43.9 μmol/L (LE = 0.36) and 7.12 μmol/L (LE = 0.42) against IDO1 and TDO, respectively (Table 2 and Fig. 3C). Since IDO1 and TDO have a characteristic absorption peak at 406 nm (Soret band) that can be as an index of the iron electronic state in heme-containing proteins (Fig. 3D)[25,30,31], we tested whether the hit compounds bind to the

heme-iron of IDO1/TDO by detecting absorption peak signals. We observed a red shift in Soret band for IDO1 and IDO incubated with **16** for 2 h at room temperature (Fig. 3D), similar as that observed for known inhibitors NLG-919 analogue and INCB-024360 (Supporting Information Fig. S11)[30], suggesting that **16** is likely to bind through the mode which involves heme-iron co-ordination probably *via* aromatic nitrogen atom, consistent with
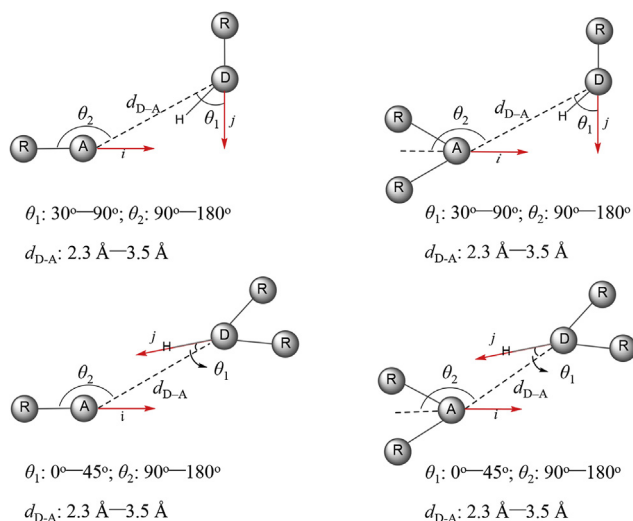
Figure 4 Definition of HA and HD. 'D' represents hydrogen bond donor, 'A' refers to hydrogen bond acceptor, 'R' is the root atom of 'D' or 'A'; d_{D-A} is the distance from 'D' to 'A', θ_1 and θ_2 represent the angles associated with root atoms. Hydrogen bonds are calculated by four categories according to the number of the root atoms.

the proposed mode by AncPhore (Fig. 3C). Similarly, **19** resulted in a red shift in Soret band when incubated with IDO1, possibly indicating that it binds to the heme-iron (Fig. 3E); **19** did not affect the UV−Vis spectrum of heme-iron in TDO, probably due to its weak TDO inhibition (Fig. 3C and E). In contrast, no red shift is observed for **20** complexed with IDO1 or TDO, possibly suggesting that **20** does not involve heme-iron coordination (Fig. 3F). Besides, **17** was also observed to affect the absorption peak signal of heme-iron in IDO1 (Fig. S11). Overall, among these hit compounds, **16** is the most promising starting point for further structural optimization to develop potential drug candidates against IDO1/TDO mediated diseases.

3. Conclusions

This work provided conceptual understanding of anchor pharmacophore features mainly in protein−ligand recognition and detailed description of AncPhore as an effective, versatile tool for drug discovery. The retrospective studies have demonstrated that consideration of the distinctive contribution and diversity of anchor pharmacophore features enables substantial improvements in the performance of AncPhore in hit/lead discovery for various types of target proteins, including metalloenzymes, which contain relatively complicate metal ion-involving interaction modes that are not well-addressed previously. The followed two case studies clearly revealed the power of AncPhore in the discovery of new structurally diverse inhibitors for MBLs and IDO1/TDO, and provided new starting points for further efforts to develop drug candidates targeting these two classes of targets. Although anchor pharmacophore features cannot be able to cover all possible binding modes, *e.g.*, allosteric effects, such previously undervalued features are an important part to improve the capacity of pharmacophore approaches in target-centered drug discovery. Moreover, understanding of the importance and diversity of pharmacophore feature contents is highly recommended, which

may be one of the inner driving forces for efficient hit/lead discovery and progressive development of pharmacophore modelling approaches.

4. Experimental

4.1. AncPhore description

AncPhore, written in C/C++ programming language, is provided as a versatile tool for pharmacophore-based drug discovery. The main features of AncPhore include pharmacophore feature analysis, and anchor pharmacophore steered molecular fitting and virtual screening. It has a large degree of flexibility to achieve different functions, such as: i) pharmacophore feature recognition and model generation for a ligand structure, an apo-protein structure or a protein−ligand complex structure; ii) ligand conformation generation by torsion-driving systematic search (TDSS)[32] and genetic algorithms (GA)[33]; iii) ligand superimposition and similarity calculation according to their pharmacophore features; iv) ligand fitting with a pharmacophore model; v) high-throughput virtual screening by single-model or multi-model manner. An additional PyMol Plugin is provided for graphic display. Detailed descriptions are as follows.

4.1.1. Definition of pharmacophore features

Ten types of pharmacophore features including HD, HA, PO, NE, MB, XB, AR, CR, HY, and CV were defined (Fig. S1). Each pharmacophore feature was labeled with a two-lettered code as shown in Supporting Information Table S4. The Gaussian volume (m, σ) was used to describe each pharmacophore feature tolerance[4], which is computed using Eq. (2):

$$V_g = \int p \exp\left(-\frac{|m-r|^2}{\sigma}\right) dr \tag{2}$$

where p is a scaling constant, m and σ represents the position and radius of each pharmacophore point, respectively. The HD, HA, MB, XB, AR, and CR features were defined with a direction vector from its center (Fig. S1).

4.1.2. Pharmacophore feature analysis for ligand−protein complex

Hydrogen bonds are one type of most common protein−ligand interactions, which occur between a hydrogen-bond donor (HD) and a hydrogen-bond acceptor (HA) in an appropriate distance and angle. Typically, the O/N/S atom with at least one hydrogen atom can be regard as an HD, and the O/N/S/F/Se atom with at least one pair of lone electrons can be regard as an HA. Herein hydrogen bonds are detected by four scenarios according to the number of the root atoms of HD and HA, and the corresponding cutoff values of θ_1, θ_2, and d_{D-A} are set[34], as shown in Fig. 4.

Electrostatic interactions between positively and negatively charged centers (PO and NE) are another common type of protein−ligand interactions. The charged centers in proteins were defined according to the electronically charged side chains of amino acids (Supporting Information Fig. S12)[35]. In ligands, the carboxylic acid, phosphoric acid, sulfonic acid, and tetrazole groups were defined as main negatively charged centers, and the ligand atoms were assigned as positively charged centers if their calculated formal charges are larger than zero. The distance cutoff
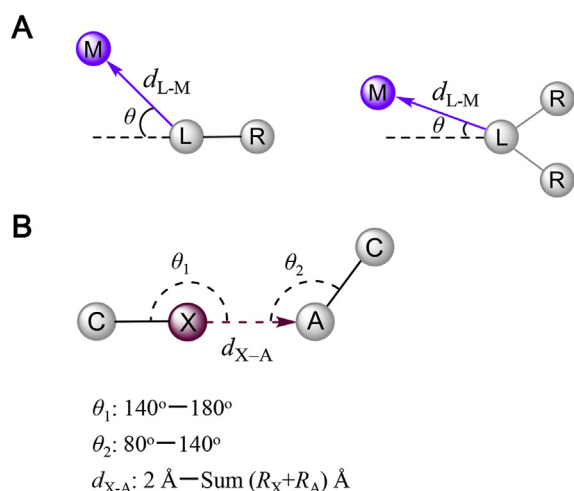
Figure 5 Definition of MB and XB. (A) 'M' is the active site metal ion(s) in the protein, 'L' refers to the ligand atom (O/N/S/F/Se) that is positioned to coordinate with metal ions, and 'R' is the root atom of 'L'; d_{L-M} is the distance from 'L' to 'M', θ is the directional bond angle, and the purple arrow represents its direction. Metal coordination bonds are calculated by two categories according to the number of 'L' root atoms. (B) 'X' is the halogen atom (I/Br/Cl) in ligand, 'A' refers to the halogen bond acceptor (O/S/N) that has at least one lone pair, and 'C' is the carbon atom. d_{X-A} is the distance from 'X' to 'A', R_X and R_A are the van der Waals radii of X and A, respectively. θ_1 is the angle of the 'A' relative to the C$-$X bond, and θ_2 is the angle of the 'X' relative to the A$-$C bond. The brown arrow represents its direction.

between a positively charged center and a negatively charged center is less than 4.5 Å.

Since metalloenzymes are widely found in various biological systems, accounting for 1/3 of all natural enzymes, metal coordination (MB) is also one type of important protein$-$ligand interactions. As observed in our previous study[11], metal coordination resembles hydrogen bonding interaction, e.g., ligand electron donor corresponding to metal ion acceptor with an obvious direction. We here defined metal coordination bonds in a similar manner as that for hydrogen bonds, which involves two parameters: d_{L-M} and θ (Fig. 5A). By analyzing metalloenzyme$-$ligand complexes, we observed that d_{L-M} and θ are commonly associated with the ligand atom types and their root atom number (Supporting Information Tables S5 and S6). Consequently, the values of ideal d_{L-M} and θ ranges are empirically derived from crystallographic structure data, as shown in Supporting Information Table S7.

Halogen bond (XB) between covalently bound halogen atoms (XB donor) and Lewis bases (XB acceptor) has important contributions to protein$-$ligand interactions[36$-$39]. Among the diverse Lewis bases, the atoms (O/S/N) possessing at least one lone pair are defined as XB acceptor in protein structures. For ligand, the heavy halogen atoms (I/Br/Cl) are considered as XB donors. XB is defined in ligand$-$protein complexes according to the distance (d_{X-A}) and the angle (θ) between XB donors and XB acceptors (Fig. 5B), where d_{X-A} is longer than 2 Å but less than the sum of the van der Waals radii of XB donor and acceptor, and θ_1 and θ_2 are in the range of 140°$-$180° and 80°$-$140°, respectively[36$-$39].

Aromatic ring (AR) involving interactions are another important type of protein$-$ligand interactions. As shown in Fig. 6A, AR

involved in $\pi-\pi$ stacking interactions are represented by the mass of the ring and the normal vector perpendicular to the ring plane, which were defined by calculating the smallest set of smallest rings and rings' aromaticity based on Huckel's rule[4,40,41]. Two kinds of typical AR interactions, i.e., face-to-face and edge-to-face interactions, are defined according to the centroid distance (d_{cen}) and the angle between ring planes (θ_1, θ_2) (Fig. 6A). The face-to-face interactions were assigned with the d_{cen} value of $\sim$4.4 Å and the θ_1 value lower than 15°, while for the edge-to-face $\pi-\pi$ interactions, d_{cen} is lower than 5.6 Å and θ_1 is in a range of 75°$-$105°, respectively[34].

Another aromatic ring involving feature is cation$-\pi$ interaction, which is emerging as one of the driving forces in molecular recognition, and is comparable to hydrogen bonding and ionic interactions[42,43]. The cation$-\pi$ interaction involves a positively charged center (PO) and an aromatic system (AR), which can be defined according to the above-described methods. Notably, metal ions in protein or ligand were also defined as cation centers. CR was finally designated according to the distance (d) and angle (θ) between the aromatic ring and the cation center in protein or ligand (Fig. 6B), where d is lower than 6.0 Å and θ is in a range of 0°$-$45°[42,43].

In order to detect the hydrophobic features (HY) in protein$-$ligand interactions, we first identify the ligand hydrophobic centers by using lipophilic contribution of ligand atoms[4]. Then, we calculate the ratio of hydrophobic residues surrounding around the hydrophobic centers within 6 Å; if the ratio exceeds an empirical cutoff value of 0.4, the hydrophobic center is defined as a hydrophobic feature along with the center position.

As the increasing number of covalent drugs and drug candidates have been reported in recent years, covalent bonding feature (CV) is considered in this study. Currently, four main types of protein residues have been found to be covalently targeted, including cysteine ($-$SH), serine/threonine/tyrosine ($-$OH), lysine ($-$NH$_2$), and glutamic acid/aspartic acid ($-$COOH). We thus annotated four modes of covalent-bonding modes, and common reactive groups for ligands[44$-$46] are shown in Supporting Information Fig. S13. The covalent bond features are defined according to the distance between protein and ligand reactive atoms without consideration of the bond direction.

4.1.3. Pharmacophore feature analysis for apo-protein structure
Here the binding site is first defined according to the co-crystal ligand or automatically identified by cavity detection. Then, a 3D grid with 0.5 Å grid space is used to cover the whole binding site. The HA, HD, PO, NE, and HY features for binding site of apo-protein are analyzed based on molecular interaction fields[35], and followed by k-means clustering analysis to yield optimal pharmacophore features[47]. Similarly, the AR features for each aromatic residue in binding site are analyzed within 3.5$-$6.1 Å and grouped into a patch which is further analyzed by clustering[47], and the direction is opposite to the normal of residues' aromatic ring. The CV feature is generated for specific residues, such as cystine, serine, threonine and lysine, if which is positioned in specific environments to form covalent bonds with small molecules. Uniquely, for MB, 16 canonical metal coordination modes[48,49], such as tetrahedron, square pyramid and octahedron, were firstly defined and established for 16 types of metal ions (Supporting Information Table S8). Then, each protein metal ion with coordination atoms (on protein) was fitted with corresponding canonical metal coordination modes; the resulted vacant positions were defined as MB features.
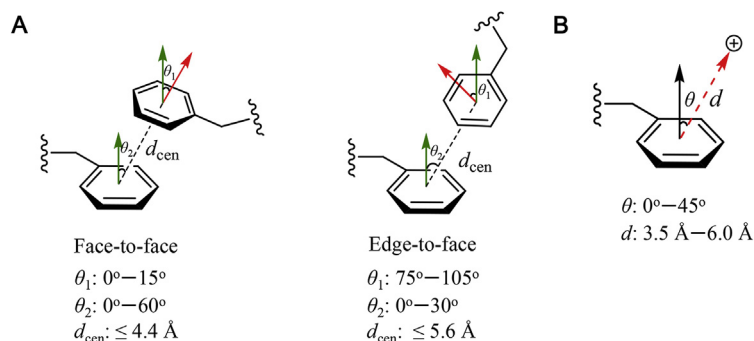
Figure 6 Definition of AR and CR. (A) An aromatic ring is represented by the mass of the ring and the normal vector perpendicular to the ring plane; AR is calculated by face-to-face and edge-to-face catalogues. (B) CR is calculated by the aromatic ring and the cation center in appropriate distance and angle.

4.1.4. Pharmacophore feature analysis for ligand

Similar as that used for protein−ligand complex, the O/N/S atom with at least one hydrogen atom was defined as an HD, and the O/N/S/F/Se atom with at least one pair of lone electrons was defined as an HA. The atom with at least one pair of lone electrons (such as O, N, S, F, and Se) with a normal vector from its root atom was defined as MB feature. The charged centers in ligands were defined according to the electronically charged atoms of groups (Fig. S12). For example, the carboxylic acid, phosphoric acid and sulfonic acid groups were defined as negatively charged centers, and the atoms were assigned as positively charged centers if their calculated formal charges are larger than zero. For AR, the center of an aromatic ring in the ligand structure was regarded as the pharmacophore point of aromatic ring with a normal vector perpendicular to the ring plane. The HY feature is defined by lipophilic contributions of ligand atoms similar as that used for pharmacophore analysis for protein−ligand complex. In addition, common reactive atoms in ligands (Fig. S13) were defined as CV. Notably, for one single atom, it may be assigned with different pharmacophore features and retained for later use in molecular fitting, virtual screening and so on.

4.1.5. Exclusion volume

The exclusion volume constraints were defined according to binding site surface atoms and mainly used to represent the shape of the binding site.

4.1.6. The protocol for cross-target pharmacophore analysis

To analyze cross-target pharmacophore features, we first classified >17,000 protein−ligand complexes collected from v2019 PDBbind database[12] into different groups by comparing protein structural similarity using the MICAN program[16]; those indexes including aligned secondary structure element length (SSEL), root mean square deviation (RMSD), and sequence identify between two proteins were considered[16]. The proteins were divided into one group if their SSEL ≥100, RMSD ≤2.0, and sequence identity ≥30%. By using this classification criteria, 201 groups were obtained; of them, 77 target protein groups contain at least three structurally similar proteins. Then, all the protein−ligand complex structures in the same group were superimposed onto one same reference complex with aim to compare them in same coordination, followed by executing pharmacophore analysis for each complex to yield a corresponding pharmacophore model. Finally, all generated pharmacophore models in one group were merged and analyzed to obtain cross-target pharmacophore features.

4.1.7. Molecular fitting

The mapping of ligand to reference pharmacophore model is achieved *via* pharmacophore superimposition between ligand pharmacophore features and reference pharmacophore features by the Largest-Overlapping-Volume-Winning[4] and total least-square algorithms, which are briefly described as follows. First, the pair of matching pharmacophore points between two set features are determined if it meets the following requirement as Eq. (3):

$$\frac{|d_{ab} - d_{cd}|}{\sigma_a + \sigma_b + \sigma_c + \sigma_d} \geq \varepsilon \tag{3}$$

where σ represents the tolerance of each pharmacophore point, d_{ab} is the distance between pharmacophore points a and b in the first pharmacophore model, d_{cd} is the distance between points c and d in the second pharmacophore model, and ε is set as 0.5. Following this requirement, all possible combinations of matching pharmacophore pairs are identified. Next, the total least-square optimization algorithm is used to determine optimal rotational angle and axis for all combinations of matching pharmacophore pairs. The total volume overlap ($V_{overlap}$) of matching pharmacophore pairs between two models is computed using Eq. (4):

$$V_{overlap} = \sum_{i=1}^{N} C_i W_i A_i \lambda_i \varphi(\theta(q)) \exp\left(\frac{-q^T P_Y q}{\sigma_{i,X} + \sigma_{i,Y}}\right) \tag{4}$$

$$\varphi(\theta(q)) = \begin{cases} \cos(\theta(q) - \theta_0) & \text{For HA, HD and MB} \\ |\cos(\theta(q))| & \text{For AR} \\ 1 & \text{For other features} \end{cases}$$

where C_i refers to a scaling factor, W_i refers to the basic weight factor reflecting inherent difference of pharmacophore features, A_i is the weight of anchor pharmacophore features, λ_i represents the weight for a chemical group in fitting with one pharmacophore feature (for example, a strong metal-binding group is treated with a higher weight than a weak metal-binding group, see Supporting Information Fig. S14), q is a unit quaternion describing rotation and translation during pharmacophore superimposition, $\theta(q)$ denotes the angle between two pharmacophore normal vectors, θ_0 is set as 0 (the number of root atoms equals 1) or $\frac{\pi}{3}$ (the number of root atoms larger than 1), $\varphi(\theta(q))$ represents the direction difference of matching pharmacophore features, and P_Y represents initial coordinates of the pharmacophore model Y. The largest $V_{overlap}$ corresponds to the best pharmacophore superimposition.

4.1.8. Scoring

The similarity value between the query and reference pharmaco-phore model is calculated using Eq. (5):

$$\text{APScore} = \frac{V_{\text{overlap}}}{V_{\text{ref}}} + k_1 \times \frac{n}{N} - k_2 \times \frac{\sum V_{\text{overlapEX}}^2}{V_{\text{refEX}}^2} \qquad (5)$$

where V_{overlap} is the largest overlapping volume between the query and reference model, V_{ref} corresponds to the volume of the reference pharmacophore model, $\sum V_{\text{overlapEX}}^2$ is the sum of square of overlapping volumes with exclusion volumes, V_{refEX}^2 corresponds to the square of exclusion volumes of the reference model, k_1/k_2 refer to the weighting factor with default values 0.2/5, n is defined as the number of matching pharmacophore pairs, and N is the number of the reference pharmacophore features. Since some specific chemical groups can be assigned with different pharma-cophore features, as, for example, a carboxyl group can represent a negatively charged center, a hydrogen bond acceptor, or a metal coordination feature, we hence use APScore as a scoring function to put particular emphasis on the reference pharmacophore model including anchor pharmacophore features, and to avoid the asymmetry of compared features, particularly in virtual screening, which will have more practicability than Tanimoto coefficient.

4.1.9. Ligand conformation generation

Another important step in AncPhore for virtual screening is to efficiently generate rational and diverse conformers for the data-base ligands of interest. The torsion-driving systematic search (TDSS) and genetic algorithm (GA) were included in AncPhore for ligand conformation generation. The TDSS method generates all diverse low-energy conformers by applying torsion-driving torsion rules and graph symmetry with energy and RMSD cutoff, as described by Oboyle et al.[32]. The GA method, mimicking the process of natural evolution, is a common method to find a global optimum solution to difficult optimization problem. It may be able to generate diverse conformers stochastically on the basis of either RMSD diversity or energy throughout gene-based evolution, e.g., mutation, crossover and replication; the computation time of conformer searching partly depends on the selection of initial population and empirical parameters.

4.1.10. Computational time

A total of 100 compounds were randomly selected for testing computational time of AncPhore on a single CPU (Intel(R) Xeon(R) CPU E5-2680, 2.40 GHz). It took about 10 and 9 min for conformation generation using TDSS and GA, respectively, and takes 25 s for these compounds in model fitting.

4.1.11. AncPhore availability

In addition to the AncPhore program, a PyMol plug-in on Linux and Window is provided for users to view the generated phar-macophore model and the computation results. The AncPhore program, PyMol plug-in and their usage instructions are freely available in https://ancphore.ddtmlab.org.

4.1.12. The strengths of AncPhore

Compared with commonly used pharmacophore tools including Pharao[4], Pharmer[5], PHASE[50], and Pharmit[51], AncPhore has its own strengths and unique characteristics, mainly including (Sup-porting Information Table S9): (1) involving MB, XB, CR, and CV features that are not included in most pharmacophore modelling tools; (2) possessing an analysis method of pharmacophore features for apo-protein, which can automatically detect cavities as possible binding sites and select corresponding protein residues for pharmacophore model generation; (3) intro-ducing a new algorithm and a new scoring function to achieve anchor pharmacophore steered molecular fitting and virtual screening; (4) supporting to use five models for molecular fitting and virtual screening, with the aim to consider pharmacophore model difference and specificity; (5) having high degree of inte-gration and portability, e.g., with two kinds of algorithms for conformation generation and convenient anchor feature assign-ment. Therefore, we believe AncPhore should be a specific and useful tool for drug discovery.

4.2. Biological assays

4.2.1. MBL protein expression and purification

The VIM-2 (residues 27−266), NDM-1 (residues 1−270), and IMP (residues 19−246) proteins were expressed and purified as described previously[52].

4.2.2. MBL inhibition assays

The experiments were performed at room temperature using Tecan microplate reader and flat-bottom 96-well black plates. Com-pounds were dissolved into 100 mmol/L with DMSO. All the MBL enzymes were diluted in the assay buffer: 20 mmol/L Tris-HCl (pH 7.5), 200 mmol/L NaCl. The compounds (with 10 different concentrations in 3-fold dilution) were pe-incubated with MBL enzymes for 10 min. The reactions were initiated by adding of the substrate FC5[53,54]. The hydrolysis of FC5 was monitored by reading the fluorescence at λ_{ex} of 380 nm and λ_{em} of 460 nm. All determinations were tested in triplicate. The IC_{50} values were calculated using GraphPad Prism software.

MBL enzymes supplemented with 1 or 100 μmol/L ZnSO_4 were pe-incubated with compounds for 10 min to examine the effects of zinc ions on compound inhibition. The reactions were initiated by the addition of FC5. The IC_{50} values were determined as described above.

4.2.3. MIC assays

The pUC57-ISAba125-pelB-VIM-2 plasmid was constructed by multi-step gene cloning and transformed into E. coli DH5α for susceptibility test. Meropenem was purchased from Dalian Meilun Biotechnology Co., Ltd. Ampicillin was obtained from Sangon Biotech (Shanghai) Co., Ltd. MHB and MHA were purchased from Solarbio (Beijing, China).

Strains of E. coli DH5α containing plasmids pUC57-ISAba125-pelB-VIM-2 (E. coli-VIM-2) and ATCC 25922 (as control) were used to assess the ability of inhibitors in restoring the antimicrobial activity of β-lactam antibiotic meropenem. Meropenem were tested alone or in combination with inhibitors at 100 or 10 μg/mL. Minimal inhibitory concentration (MIC) values were determined by the standard broth micro-dilution method according to the Clinical and Laboratory Standards Institute (CLSI, M07-A9, 2012) guideline.

4.2.4. Crystallization and data collection for the VIM-2:4 complex

The VIM-2:4 complexes were crystallized by hanging drop vapor diffusion method as described previously[52]. Briefly, 10 mg/mL VIM-2 proteins and 5 mmol/L 4 were incubated in crystallization buffer (20 mmol/L Tris-HCl, pH 7.5, 200 mmol/L NaCl, 0.5 mmol/L tris (2-carboxyethyl) phosphine (TCEP) for 60 min at

4 °C. One microliter (μL) of protein-inhibitor solution was mixed with 1 μL of precipitant (25%−32% PEG 3350, 0.2 mol/L magnesium formate) and incubated at 20 °C. The crystals were harvested using the cryo-protectant solution 30% (v/v) glycerol, and flash-cooled in liquid nitrogen. The diffraction data were collected at the Shanghai Synchrotron Radiation Facility, and processed using HKL2000. The structure was solved by molecular replacement and refinement using Phenix[55] and Coot[56] programs. Coordinates and structure factors of the VIM-2:**4** complex structure (PDB code 7CHV) have been deposited in the PDB.

4.2.5.   IDO1/TDO protein expression and purification

The human IDO1 (residues 12−403) and TDO (residues 19−388)[26−28] were cloned into pET28a vectors for expression with N-terminally His$_6$-Tagged proteins. IDO1/TDO were over-expressed in *E. coli* Transetta (DE3) cells at 37 °C using LB medium supplied with 1 mmol/L 5-ALA and 30 μmol/L hemin chloride in a shaker at 200 rpm. When the OD$_{600}$ value reached 0.8−1.0, the temperature was lowered to 25 °C, followed by addition of 0.5 mmol/L IPTG to the culture to induce protein expression for 6−8 h in a shaker at 150 rpm. Cells were then harvested by centrifugation (20 min, 4000 rpm), resuspended in lysis buffer A (for IDO1: 50 mmol/L potassium phosphate, pH 7.1, 0.3 mol/L sodium chloride, 25 mmol/L imidazole, 5% glycerol; for TDO: 50 mmol/L potassium phosphate, pH 7.8, 0.3 mol/L sodium chloride, 5% glycerol) supplemented with EDTA-free protease inhibitor, and lysed by using an ultrahigh-pressure homogenizer (JNBIO). The lysates were clarified by sedimentation (13,000 rpm) for 30 min at 4 °C, and then loaded onto an Ni-NTA column, followed by extensive washing with buffer B (for IDO1: 50 mmol/L potassium phosphate, pH 7.1, 0.3 mol/L sodium chloride, 30 mmol/L imidazole, 5% glycerol; for TDO: 50 mmol/L potassium phosphate, pH 7.8, 0.3 mol/L sodium chloride, 30 mmol/L imidazole, 5% glycerol) to remove nonspecifically binding proteins. The target proteins were eluted with buffer C (for IDO1: 50 mmol/L potassium phosphate, pH 7.1, 0.3 mol/L sodium chloride, 350 mmol/L imidazole, 5% glycerol; for TDO: 50 mmol/L potassium phosphate, pH 7.8, 0.3 mol/L sodium chloride, 250 mmol/L imidazole, 5% glycerol). Fractions containing the purified enzymes were concentrated using Amicon Ultra 10K (Millipore) and then desalted using a HiTrap desalting column (GE Healthcare) into pH 6.5 potassium phosphate buffer (100 mmol/L). The purified proteins were stored at −80 °C before use.

4.2.6.   IDO1/TDO inhibition assays

The IDO1/TDO enzymatic activities were tested in the assay buffer: 100 mmol/L potassium phosphate, pH 6.5, 0.01% Triton X-100, 40 mmol/L ascorbic acid, 7 μmol/L methylene blue, and 200 μg/mL catalase. The compounds (10 different concentrations) were pre-incubated with the enzymes for 5 min at 25 °C. The reactions were initiated by addition of L-trp (400 μmol/L) in the assay buffer and incubated 30−60 min at 37 °C. The reactions were stopped by addition of 30 μL 30% TCA and incubated 30 min at 50 °C to convert *N*-formylkynurenine produced by IDO1/TDO into kynurenine. Then, 90 μL of 2% (w/v) pDMAB in acetic acid were added and incubated for 5 min at room temperature. The absorbance at 480 nm was measured using Tecan microplate reader. All determinations were tested in triplicate.

4.2.7.   Optical absorption spectroscopic measurements

Ultraviolet−visible (UV−Vis) absorption spectra of IDO1/TDO in complex with various inhibitors were recorded at room temperature under standard atmosphere using Varioskan LUX multifunctional reader (Thermo Fisher Scientific Co., Ltd.) with a spectral slit width of 1 nm. The samples were prepared with 10 μmol/L ferric IDO1 or TDO and 0.5 or 1.0 mmol/L inhibitors in 100 mmol/L potassium phosphate buffer (pH 6.5). UV−Vis absorption spectra (300−700 nm) were recorded after samples incubated for 2 h at room temperature.

Author contributions

Qingqing Dai and Yuhang Yan contributed equally. Guo-Bo Li designed this project. Qingqing Dai wrote the programming codes; Qingqing Dai, Gen Li, and Junlin Yu tested the programs. Yuhang Yan, Xiangli Ning, Ji Deng, and Lingling Yang performed the biological tests. Guo-Bo Li and Qingqing Dai wrote the manuscript. All authors have given approval to the final version of the manuscript.

Conflicts of interest

The authors declare no competing financial interest.

Appendix A.   Supporting information

Supporting data to this article can be found online at https://doi. org/10.1016/j.apsb.2021.01.018.

References

1. Du J, Guo J, Kang D, Li Z, Wang G, Wu J, et al. New techniques and strategies in drug discovery. *Chin Chem Lett* 2020;**31**:1695−708.
2. Yang SY. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discov Today* 2010;**15**: 444−50.
3. Schaller D, Šribar D, Noonan T, Deng L, Nguyen TN, Pach S, et al. Next generation 3D pharmacophore modeling. *WIREs Comput Mol Sci* 2020;**10**:e1468.
4. Taminau J, Thijs G, De Winter H. Pharao: Pharmacophore alignment and optimization. *J Mol Graph Model* 2008;**27**:161−9.
5. Koes DR, Camacho CJ. Pharmer: Efficient and exact pharmacophore search. *J Chem Inf Model* 2011;**51**:1307−14.
6. Roskoski R. Properties of FDA-approved small molecule protein kinase inhibitors: a 2020 update. *Pharmacol Res* 2020;**152**:104609.
7. Yan YH, Li G, Li GB. Principles and current strategies targeting metallo-β-lactamase mediated antibacterial resistance. *Med Res Rev* 2020;**40**:1558−92.
8. Wang YL, Liu S, Yu ZJ, Lei Y, Huang MY, Yan YH, et al. Structure-based development of (1-(3′-mercaptopropanamido)methyl)boronic

acid derived broad-spectrum, dual-action inhibitors of metallo- and serine-β-lactamases. *J Med Chem* 2019;**62**:7160−84.

9. Li GB, Abboud MI, Brem J, Someya H, Lohans CT, Yang S, et al. NMR-filtered virtual screening leads to non-metal chelating metallo-β-lactamase inhibitors. *Chem Sci* 2017;**8**:928−37.

10. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. *J Med Chem* 2012;**55**:6582−94.

11. Li G, Su Y, Yan YH, Peng JY, Dai QQ, Ning XL, et al. MeLAD: an integrated resource for metalloenzyme-ligand associations. *Bioinformatics* 2020;**36**:904−9.

12. Liu Z, Li Y, Han L, Li J, Liu J, Zhao Z, et al. PDB-wide collection of binding data: current status of the PDBbind database. *Bioinformatics* 2014;**31**:405−12.

13. Liu Z, Su M, Han L, Liu J, Yang Q, Li Y, et al. Forging the basis for developing protein−ligand interaction scoring functions. *Acc Chem Res* 2017;**50**:302−9.

14. Wang R, Fang X, Lu Y, Yang CY, Wang S. The PDBbind database: methodologies and updates. *J Med Chem* 2005;**48**:4111−9.

15. Alterio V, Di Fiore A, D'Ambrosio K, Supuran CT, De Simone G. Multiple binding modes of inhibitors to carbonic anhydrases: how to design specific drugs targeting 15 different isoforms. *Chem Rev* 2012;**112**:4421−68.

16. Minami S, Sawada K, Chikenji G. Mican: a protein structure alignment algorithm that can handle multiple-chains, inverse alignments, Cα only models, alternative alignments, and non-sequential alignments. *BMC Bioinform* 2013;**14**:24.

17. Prabu-Jeyabalan M, Nalivaika E, Schiffer CA. How does a symmetric dimer recognize an asymmetric substrate? A substrate complex of HIV-1 protease. *J Mol Biol* 2000;**301**:1207−20.

18. Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* 2006;**125**:1137−49.

19. Zheng J, Trafny EA, Knighton DR, Xuong N, Taylor SS, Ten Eyck LF, et al. 2.2 Å refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. *Acta Crystallogr D* 1993;**49**:362−5.

20. Cukuroglu E, Engin HB, Gursoy A, Keskin O. Hot spots in protein−protein interfaces: towards drug discovery. *Prog Biophys Mol Biol* 2014;**116**:165−73.

21. Salimraj R, Hinchliffe P, Kosmopoulou M, Tyrrell JM, Brem J, van Berkel SS, et al. Crystal structures of VIM-1 complexes explain active site heterogeneity in VIM-class metallo-β-lactamases. *FEBS J* 2019;**286**:169−83.

22. Li GB, Brem J, Lesniak R, Abboud MI, Lohans CT, Clifton IJ, et al. Crystallographic analyses of isoquinoline complexes reveal a new mode of metallo-β-lactamase inhibition. *Chem Commun* 2017;**53**:5806−9.

23. Vécsei L, Szalárdy L, Fülöp F, Toldi J. Kynurenines in the CNS: recent advances and new questions. *Nat Rev Drug Discov* 2012;**12**:64.

24. Liang H, Chen M, Qi F, Shi L, Duan Z, Yang R, et al. The proatherosclerotic function of indoleamine 2,3-dioxygenase 1 in the developmental stage of atherosclerosis. *Signal Transduct Target Ther* 2019;**4**:23.

25. Röhrig UF, Reynaud A, Majjigapu SR, Vogel P, Pojer F, Zoete V. Inhibition mechanisms of indoleamine 2,3-dioxygenase 1 (IDO1). *J Med Chem* 2019;**62**:8784−95.

26. Yang D, Zhang S, Fang X, Guo L, Hu N, Guo Z, et al. *N*-benzyl/aryl substituted tryptanthrin as dual inhibitors of indoleamine 2,3-dioxygenase and tryptophan 2,3-dioxygenase. *J Med Chem* 2019;**62**:9161−74.

27. Serafini M, Torre E, Aprile S, Grosso ED, Gesù A, Griglio A, et al. Discovery of highly potent benzimidazole derivatives as indoleamine 2,3-dioxygenase-1 (IDO1) inhibitors: from structure-based virtual screening to in vivo pharmacodynamic activity. *J Med Chem* 2020;**63**:3047−65.

28. Kazmierski WM, Xia B, Miller J, De la Rosa M, Favre D, Dunham RM, et al. DNA-encoded library technology-based discovery, lead optimization, and prodrug strategy toward structurally unique

indoleamine 2,3-dioxygenase-1 (IDO1) Inhibitors. *J Med Chem* 2020;**63**:3552−62.

29. Kumar S, Waldo JP, Jaipuri FA, Marcinowicz A, Van Allen C, Adams J, et al. Discovery of clinical candidate (1*R*,4*r*)-4-((*R*)-2-((*S*)-6-fluoro-5*H*-imidazo[5,1-*a*]isoindol-5-yl)-1-hydroxyethyl)cyclohexan-1-ol (navoximod), a potent and selective inhibitor of indoleamine 2,3-dioxygenase 1. *J Med Chem* 2019;**62**:6705−33.

30. Lewis-Ballester A, Pham KN, Batabyal D, Karkashon S, Bonanno JB, Poulos TL, et al. Structural insights into substrate and inhibitor binding sites in human indoleamine 2,3-dioxygenase 1. *Nat Commun* 2017;**8**:1693.

31. Pham KN, Lewis-Ballester A, Yeh SR. Structural basis of inhibitor selectivity in human indoleamine 2,3-dioxygenase 1 and tryptophan dioxygenase. *J Am Chem Soc* 2019;**141**:18771−9.

32. Oboyle NM, Vandermeersch T, Flynn CJ, Maguire AR, Hutchison GR. Confab systematic generation of diverse low-energy conformers. *J Cheminformatics* 2011;**3**:8.

33. Judson RS, Jaeger EP, Treasurywala AM, Peterson ML. Conformational searching methods for small molecules. II. Genetic algorithm approach. *J Comput Chem* 1993;**14**:1407−14.

34. Wang Z, Ji H. Targeting the side-chain convergence of hydrophobic α-helical hot spots to design small-molecule mimetics: key binding features for i, i + 3, and i + 7. *J Med Chem* 2019;**62**:9906−17.

35. Li GB, Yang LL, Wang WJ, Li LL, Yang SY. ID-Score: a new empirical scoring function based on a comprehensive set of descriptors related to protein−ligand interactions. *J Chem Inf Model* 2013;**53**:592−600.

36. Cavallo G, Metrangolo P, Milani R, Pilati T, Priimagi A, Resnati G, et al. The halogen bond. *Chem Rev* 2016;**116**:2478−601.

37. Shinada NK, de Brevern AG, Schmidtke P. Halogens in protein−ligand binding mechanism: a structural perspective. *J Med Chem* 2019;**62**:9341−56.

38. Auffinger P, Hays FA, Westhof E, Ho PS. Halogen bonds in biological molecules. *Proc Natl Acad Sci U S A* 2004;**101**:16789−94.

39. Xu Z, Yang Z, Liu Y, Lu Y, Chen K, Zhu W. Halogen bond: its role beyond drug−target binding affinity for drug discovery and development. *J Chem Inf Model* 2014;**54**:69−78.

40. Figueras J. Ring perception using breadth-first search. *J Chem Inf Comput Sci* 1996;**36**:986−91.

41. Roos-Kozel BL, Jorgensen WL. Computer-assisted mechanistic evaluation of organic reactions. 2. Perception of rings, aromaticity, and tautomers. *J Chem Inf Comput Sci* 1981;**21**:101−11.

42. Minoux H, Chipot C. Cation−π interactions in proteins: can simple models provide an accurate description. *J Am Chem Soc* 1999;**121**:10366−72.

43. Dougherty DA. The cation−π interaction. *Acc Chem Res* 2013;**46**:885−93.

44. Ghosh AK, Samanta I, Mondal A, Liu WR. Covalent inhibition in drug discovery. *ChemMedChem* 2019;**14**:889−906.

45. Gehringer M, Laufer SA. Emerging and re-emerging warheads for targeted covalent inhibitors: applications in medicinal chemistry and chemical biology. *J Med Chem* 2019;**62**:5673−724.

46. Yang B, Wang WJ, Li LL. Research progress of the small molecule covalent inhibitors. *Acta Pharm Sin* 2014;**49**:158−65.

47. Hu B, Lill MA. Exploring the potential of protein-based pharmacophore models in ligand pose prediction and ranking. *J Chem Inf Model* 2013;**53**:1179−90.

48. Putignano V, Rosato A, Banci L, Andreini C. MetalPDB in 2018: a database of metal sites in biological macromolecular structures. *Nucleic Acids Res* 2018;**46**:D459−64.

49. Andreini C, Cavallaro G, Lorenzini S, Rosato A. MetalPDB: a database of metal sites in biological macromolecular structures. *Nucleic Acids Res* 2013;**41**:D312−9.

50. Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. Phase: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J Comput Aided Mol Des* 2006;**20**:647−71.

51. Sunseri J, Koes DR. Pharmit: interactive exploration of chemical space. *Nucleic Acids Res* 2016;**44**:W442−8.

52. Wang YL, Liu S, Yu ZJ, Lei Y, Huang MY, Yan YH, et al. Structure-based development of (1-(3′-mercaptopropanamido)methyl) boronic acid derived broad-spectrum, dual-action inhibitors of metallo- and serine-beta-lactamases. *J Med Chem* 2019;**62**:7160−84.

53. van Berkel SS, Brem J, Rydzik AM, Salimraj R, Cain R, Verma A, et al. Assay platform for clinically relevant metallo-*β*-lactamases. *J Med Chem* 2013;**56**:6945−53.

54. Liu S, Jing L, Yu ZJ, Wu C, Zheng Y, Zhang E, et al. ((*S*)-3-Mercapto-2-methylpropanamido)acetic acid derivatives as metallo-*β*-lactamase inhibitors: synthesis, kinetic and crystallographic studies. *Eur J Med Chem* 2018;**145**:649−60.

55. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. Phenix: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 2010;**66**:213−21.

56. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004;**60**:2126−32.