RESEARCH ARTICLE

[View Article Online](#)
[View Journal](#) | [View Issue](#) Cite this: *Med. Chem. Commun.*,
2018, 9, 937

Discovery of indoleamine 2,3-dioxygenase inhibitors using machine learning based virtual screening†

Hongao Zhang, Wei Liu, Zhihong Liu, Yingchen Ju, Mengyang Xu, Yue Zhang, Xinyu Wu, Qiong Gu, * Zhong Wang* and Jun Xu *

Indoleamine 2,3-dioxygenase (IDO), an immune checkpoint, is a promising target for cancer immunotherapy. However, current IDO inhibitors are not approved for clinical use yet; therefore, new IDO inhibitors are still demanded. To identify new IDO inhibitors, we have built naive Bayesian (NB) and recursive partitioning (RP) models from a library of known IDO inhibitors derived from recent publications. Thirteen molecular fingerprints were used as descriptors for the models to predict IDO inhibitors. An in-house compound library was virtually screened using the best machine learning model, which resulted in 50 hits for further enzyme-based IDO inhibitory assays. Consequently, we identified three new IDO inhibitors with IC_{50} values of 1.30, 4.10, and 4.68 μ M. These active compounds also showed IDO inhibitory activities in cell-based assays. The compounds belong to the tanshinone family, a typical scaffold family derived from Danshen (a Chinese herb), the dried root of *Salvia miltiorrhiza*, which has been widely used in China, Japan, the United States, and other European countries for the treatment of cardiovascular and cerebrovascular diseases. Thus, we discovered a new use for Danshen using machine learning methods. Surface plasmon resonance (SPR) experiments proved that the inhibitors interacted with the IDO target. Molecular dynamic simulations demonstrated the binding modes of the IDO inhibitors.

Received 21st December 2017,
Accepted 24th February 2018

DOI: 10.1039/c7md00642j

rsc.li/medchemcomm

1. Introduction

Indoleamine 2,3-dioxygenase (IDO) is an intracellular heme-containing enzyme, which controls the rate-limiting steps in the metabolism of tryptophan (Trp) along the kynurenine (Kyn) pathway. The IDO family has two homologs, IDO1 and IDO2. IDO2 appears to be less efficient, and no ideal assay conditions *in vitro* for IDO2 were reported. Therefore, IDO means IDO1 in this paper.¹

IDO is expressed in many tissues and cell types, such as in the placenta, lungs, small and large intestines, colon, spleen, liver, kidneys, stomach, brain, tumor cell lines, dendritic cells, and macrophages.^{2,3} It is associated with several diseases, such as infectious diseases, cancer, neurological diseases, and autoimmune diseases such as type 1 diabetes. For infectious diseases, a pathogenic microorganism can induce interferon- γ (IFN- γ), nuclear factor κ B (NF- κ B), and IDO gene

overexpression to metabolize more Trp to prevent Trp from being consumed by pathogenic microorganisms.⁴ T cells are also impressible to a short-of-Trp microenvironment. Under the influence of immunosuppressive antigen presenting cells induced by IDO, T cells cease proliferation and alter immunosuppressive regulatory T cells. Immunosuppression is also in favor of infection. Hence, IDO plays two contrary roles in infectious diseases.^{5–7} Cancer cells, however, utilize the immunosuppressive mechanism to escape clearance of the immune system by mutations. Experiments suggest that IDO can affect the downstream signaling pathways of other immune checkpoints, PD-1 and CTLA-4.^{8–10} In neurological diseases, decreasing Trp will affect serotonin synthesis and cause depression.¹¹ Trp metabolites in the Kyn pathway are injurious to an organism, and are associated with Alzheimer's disease (AD), Parkinson's disease (PD), and Huntington's disease (HD).^{12–14} Animal experiments indicate that IDO inhibition can be a therapeutic strategy against neoplastic and neurologic diseases.^{15–18}

Several IDO inhibitors entered clinical trials, none of them were approved by the US Food and Drug Administration (FDA) yet. Hence, the discovery of new IDO inhibitors is still demanded.^{19–22}

Research Center for Drug Discovery, School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, China. E-mail: guqiong@mail.sysu.edu.cn, wangzh357@mail.sysu.edu.cn, junxu@biochemomes.com

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7md00642j

2. Results and discussion

2.1. Diversity analysis of the dataset

The predictive ability of a model depends on the structural diversity of the training compounds. As shown in Fig. 1A, our training compounds are structurally diversified. The distributions of other molecular properties also indicate that our

compounds are drug-like. The p -values of all the properties including $A \log P$ (logarithm partition coefficient), $\log D$ (logarithm distribution coefficient), MW (molecular weight), solubility, pK_a , HBA (hydrogen bond acceptors), HBD (hydrogen bond donors), N-counts (nitrogen atom counts), and O-counts (oxygen atom counts) are depicted in Fig. 1B, which indicates that MW, HBA, and O-counts are descriptors that

Fig. 1 Diversity analysis of the dataset. (A) SCA plot for the structural diversity of the compounds; (B) the distribution of properties of active compounds and inactive compounds.

can discriminate actives from inactives. The correlation coefficient for descriptors O-counts and pIC_{50} is 0.399 (Table S1†), suggesting that the two descriptors are more correlated than the other descriptors.

2.2. Performance of the models

86 RP models and 86 NB models were generated. The number of models having Q (overall predictive accuracy) values greater than 0.8 is 69. As shown in Fig. 2A, the Matthews

Fig. 2 (A) The relations between MCC and fingerprint diameters. (B) The relations between Q /MCC/AUC and fingerprint types for the models with training and testing data.

correlation coefficient (MCC) values of the fingerprints with different diameters are fluctuating. The average MCC values for the training set and the test set are greater than 0.74 and 0.54 if the diameter is greater than 4. For the test sets (Fig. 2B), the *Q* values of the top-10 models are greater than 0.76, the MCC values are greater than 0.53, and the values of the area under the receiver operating characteristic curve (AUC) are greater than 0.89. This suggests that the models are of good quality.

2.3. Validating the models with external testing data

The 4 RP models manifested better performance than the 6 NB models (Table 1). The 4 RP models were based on fingerprints FPFC2, FFPF6, FPFC8 and FFPF10 with MCC values of 0.487, 0.430, 0.458 and 0.458, respectively, and AUC values of 0.813, 0.791, 0.815 and 0.792, respectively.

2.4. Virtual screening using selected models

The RP model with fingerprint FPFC8 is selected as the winning model for virtually screening our in-house compound library (8333 compounds) for IDO inhibitors. The screening campaign resulted in 50 virtual hits after removing the compounds with known IDO inhibiting scaffolds. The main scaffolds of the hit compounds are tanshinone, flavone, quinolone, and sulfanilamide.

2.5. IDO inhibitory assay

The IDO inhibitor NLG8189 was used as the positive control agent. The IC_{50} value of NLG8189 is 696.7 μ M. All the hits were tested at 10 μ M. Some compounds were not tested for IDO inhibitory activity due to their potential fluorescence properties. The detailed information about the hits can be found in the ESI† (Table S2). Further experiments demonstrated that compounds SYSU-00440, SYSU-00464, and SYSU-00469 are IDO inhibitors with IC_{50} values of 1.30, 4.10, and 4.68 μ M (Table 2). Compounds SYSU-00440 and SYSU-00469 showed higher inhibition in cell-based assays (Table 2).

Table 1 The performance of the top-10 models with external testing data^a

RP	Test set	TP	TN	FP	FN	SE	SP	Q	MCC	AUC
	FPFC2	49	26	5	22	0.690	0.839	0.735	0.487	0.813
	FFPF6	53	22	9	18	0.746	0.710	0.735	0.430	0.791
	FPFC8	53	23	8	18	0.746	0.742	0.745	0.458	0.815
	FFPF10	53	23	8	18	0.746	0.742	0.745	0.458	0.792
NB	Test set	TP	TN	FP	FN	SE	SP	Q	MCC	AUC
	FCFP6	8	31	0	63	0.113	1.000	0.382	0.193	0.704
	ECFP8	25	27	4	46	0.352	0.871	0.510	0.227	0.737
	FCFP8	9	31	0	62	0.127	1.000	0.392	0.206	0.705
	ECFP10	26	27	4	45	0.366	0.871	0.520	0.239	0.739
	ECFP12	35	25	6	36	0.493	0.806	0.588	0.281	0.739
	FCFP12	33	26	5	38	0.465	0.839	0.578	0.289	0.695

^a TP: true positives, TN: true negatives, FP: false positives, FN: false negatives, SE: sensitivity, SP: specificity.

2.6. SPR results

To confirm that compounds SYSU-00440, SYSU-00464, and SYSU-00469 bind to IDO, SPR experiments were conducted. The representative binding sensorgram (for the interactions of SYSU-00469 and IDO) is depicted in Fig. 3. Compounds SYSU-00440 and SYSU-00464 demonstrated similar binding behavior. The sensorgrams for SYSU-00440 and SYSU-00464 binding to IDO are depicted in the ESI† Fig. S3.

2.7. Binding modes of the confirmed hits

Compounds SYSU-00440, SYSU-00464, and SYSU-00469 were docked into the binding pocket of an IDO crystal structure (PDB access code: 5EK3). Molecular dynamic simulations demonstrate that the three compounds bind to Fe^{2+} of heme, and interact with residues Y126, F163, and F226 within hydrophobic pockets (Fig. 4).²² The compounds also interact with R231 at the entrance to the pocket and occupy the catalytic site. The estimated ΔG values for the compounds (SYSU-00440, SYSU-00464, SYSU-00469) and IDO complexes were -18.80 ± 2.68 , -10.59 ± 3.42 , and -14.92 ± 3.23 kcal mol⁻¹. These ΔG values are consistent with the IC_{50} values of the three active compounds.

2.8. The new use of tanshinone derivatives

As IDO inhibitors, compounds SYSU-00440, SYSU-00464, and SYSU-00469 are tanshinones derived from Danshen, the dried root of *Salvia miltiorrhiza*. Danshen, a Chinese herb, has been used in China, Japan, the United States, and other European countries for the treatment of cardiovascular and cerebrovascular diseases.⁴⁶ In China, 38 approved traditional Chinese medicines (TCM) contain tanshinones. Tanshinones can induce the maturation of human dendritic cells through the activation of NF- κ B, p38 and JNK MAPKs, and may have the capacity to promote adaptive immune responses.⁴⁷ Cryptotanshinone induces ROS-dependent autophagy in multidrug-resistant colon cancer cells.⁴⁸ Our studies demonstrate that tanshinones can be repurposed for anti-cancer treatments.

3. Conclusions

IDO is a promising target for cancer therapy. This study has demonstrated that machine learning methods are capable of identifying IDO inhibitors with new scaffolds. Among the new IDO inhibitors, three are regarded as active agents (IC_{50} less than 5 μ M). Specifically, compounds SYSU-00440 and SYSU-00469 showed potent inhibitory activity in cell-based assays.

4. Materials and methods

High-throughput screening (HTS) and virtual screening approaches^{23,24} are combined for lead identification in this work. Naive Bayesian (NB) and recursive partitioning (RP) learning methods were employed to build 172 models using

Table 2 The IDO inhibitory activity of the hits in enzyme- and cell-based assays

Enzyme experiment		Cell experiment		
Compound	IC ₅₀ (μM)	Concentration (μM)	Survival rate (%)	Percentage of inhibition (%)
SYSU-00440	1.30 ± 0.28	1	94.72 ± 5.82	54.34 ± 3.17

SYSU-00464	4.10 ± 0.32	10	97.13 ± 4.11	20.97 ± 2.97

SYSU-00469	4.68 ± 0.08	1	93.00 ± 3.70	66.66 ± 1.18

NLG8189 ^a	696.7 ± 1.5			

^a NLG8189 was used as the positive control. **Fig. 3** Sensorgram for the interactions of SYSU-00469 with IDO.

the data set of 504 compounds (242 active compounds and 262 inactive compounds), with data on inhibitory activity against IDO. The models were validated with a 5-fold cross validation method. The best model was selected by an external testing process. An in-house compound library, which had 8333 tangible compounds (5772 synthetic compounds and 2561 natural compounds), was virtually screened using the best model. Fifty hit compounds were further validated *via* enzyme-based and cell-based assays, and the binding abilities and binding modes of potent inhibitors were experimentally confirmed. The flow chart of the process is depicted in Fig. 5.

4.1. Data collection

The IDO inhibitors used in this work were derived from the ChEMBL database.²⁵ The recently published data of IDO inhibitors were collected from the literature.^{22,26–30} All the data were checked and selected using the following criteria: (1) the enzyme-based inhibitory assay with Kyn detection as the

index; (2) the compound having an IC₅₀ value for IDO; (3) removing duplicated compounds and cell assay data; (4) IC₅₀ values and structures of the compounds being checked against original references. The activity threshold was set to 10 μM. The 402 compounds derived from the ChEMBL database were divided into training sets (302 compounds) and test sets (100 compounds) randomly. The 102 compounds from recent studies were collected for external testing or validation.

4.2. Diversity analysis

The structural diversity of the chemical structures of the data library was analyzed using the S-cluster approach.³¹ Cyclicity and complexity are two indexes measuring the structural diversity of the library. The molecular properties were calculated using DS 3.5 (Discovery Studio version 3.5, Accelrys Inc., USA). Student's *t*-test was used to calculate the *p*-value of all the properties of active and inactive compounds.³² The *p*-value and the correlation coefficients between the

Fig. 4 Binding modes of the confirmed hits. (A)–(C) for SYSU-00440, SYSU-00464, SYSU-00496 respectively. Hydrogen bonds are labeled as orange dashed lines, and coordination bonds are labeled as violet dashed lines.

properties and inhibitory activity were calculated using SPSS 13.0. The SCA plot and diversity map were also used to pick compounds showing a higher possibility of inhibitory activity.

4.3. Descriptors

Structural fingerprints were calculated and utilized as descriptors for the QSAR studies in this paper. The selected structural fingerprints were ECFC, ECFP, EPFC, EPFP, FCFC, FCFP, FPFC, FPFP, LCFC, LCFP, LPFC, LPFP, and SEFP with diameters of 0, 2, 4, 6, 8, 10, and 12, which were calculated using DS 3.5 software.

4.4. Modeling methods

Naive Bayesian (NB) and recursive partitioning (RP) machine learning models were constructed using DS 3.5 software.

Fig. 5 The flow chart of the IDO inhibitor discovery process.

4.4.1. Naive Bayesian (NB). The NB method is a probability classifier based on the Bayesian principle and the maximum posteriori hypothesis. Based on the prior distribution of parameters, the overall distribution is calculated directly, and the new sample data will be predicted depending on the overall distribution of parameters.^{33,34}

4.4.2. Recursive partitioning (RP). The RP method recursively divides a data set into smaller subsets, which results in a hierarchical tree (aka decision tree), which represents relationships among data points and independent descriptors.³⁵ In this study, the tree depth was set to 20.

4.5. Evaluating the models

The performances of the models were evaluated with a 5-fold cross validation process. True positives (TP), true negatives (TN), false positives (FP), false negatives (FN), the sensitivity (SE), the specificity (SP), the overall predictive accuracy (Q), the Matthews correlation coefficient (MCC), and the values of the area under the receiver operating characteristic curve (AUC) were calculated.³⁶ Moreover, an independent external test set was used to evaluate the actual prediction power of the models.

4.6. IDO protein expression and purification

The DNA fragment (reverse transcribed from RNA which was extracted from Huh-7 cells) encoding for the IDO was generated by PCR and inserted into the NdeI and XhoI sites of the pET-15b vector (Novagen). The *E. coli* strain BL21 codon plus (Invitrogen) cells carried the IDO plasmids and were grown at 37 °C overnight. The overnight culture was seeded into fresh

medium and incubated at 37 °C with shaking until the culture reached an OD₆₀₀ of 0.4. It was induced with 0.1 mM isopropyl- β -D-thio-galactoside (IPTG) for 16 h at 20 °C. The cells were harvested by centrifugation at 4000 rpm for 30 min at 4 °C and resuspended with binding buffer. After sonication, the cell lysates were centrifuged at 18000 rpm for 30 min, and the supernatant was loaded in a Ni-NTA column (QIAGEN).³⁷ The IDO protein was purified by size-exclusion chromatography (GE Healthcare, Hiload Superdex 200 16/60). The purified IDO proteins were desalted and concentrated to 30 mg mL⁻¹ in storage buffer and stored at -80 °C.

4.7. IDO enzyme assay

The IDO inhibitory activity assay was carried out as described in the reference with some modifications.³⁸ In brief, 90 mL of the reaction mixture in a 96-well black plate contained 50 mM potassium phosphate buffer (pH 6.5), 20 mM ascorbic acid, 10 μ M methylene blue, 100 μ g mL⁻¹ catalase, 200 μ M L-tryptophan and the test compounds (10 mM in DMSO diluted with 50 mM potassium phosphate buffer before use). The fluorescence (λ_{ex} 360 nm, λ_{em} 480 nm) was measured. The plate was incubated at 37 °C for 1 h after 10 μ L of 20 μ g mL⁻¹ IDO was added. To stop the reaction, 20 μ L of 1 M NaOH solution was added and the plate was incubated at 60 °C for 15 min. The plate was placed at room temperature for about 1.5 h before the measurement of fluorescence (λ_{ex} 360 nm, λ_{em} 480 nm). The percentage of inhibition was reported as $(100 - (A/B \times 100))$, where A is the IDO activity in the presence of the inhibitor and B is the IDO activity in the absence of the inhibitor. The IC₅₀ values were generated from 3 independent experiments with different concentration points. The results were calculated using GraphPad Prism 7.

4.8. Cell-based IDO assay

The Huh-7 cell line (cells were provided by the Stem Cell Bank, Chinese Academy of Sciences) was used in the cell-based IDO assay.³⁹ After the cells were seeded into the 24-well plate for 24 h, 500 μ L culture medium (DMEM with 10% FBS, 1% penicillin-streptomycin, 70 μ M L-tryptophan, 50 ng mL⁻¹ IFN- γ , 10 μ M or 1 μ M compound) was added, then the previous medium was discarded, and cultivated in an incubator for 24 h. DMEM and FBS were purchased from Thermo Fisher, and IFN- γ was purchased from Sangon Biotech. The supernatant was transferred to 1.5 mL tubes and mixed with 100 μ L 50% (w/v) trichloroacetic acid (TCA). The tubes were incubated at 65 °C in a water bath for 30 min, followed by centrifugation at 12000 rpm for 10 min. The clear supernatant (125 μ L) was transferred to a new flat-bottomed 96-well plate and mixed with an equal volume of 2% (w/v) *p*-dimethylamino-benzaldehyde in acetic acid. The resulting reaction was measured at OD 490 nm.⁴⁰ Three independent assays were conducted to determine the inhibition of every compound. The survival rate of the cell was tested by MTT assays under the above-mentioned conditions.⁴¹

4.9. Surface plasmon resonance (SPR) experiments

SPR experiments were carried out using a ProteOn XPR36TM SPR instrument (BioRad Hercules, CA). Standard amine coupling was used to immobilize IDO (20 nM in a 10 mM sodium acetate buffer, pH 5.5) on the EDAC/sulfo-NHS activated surface of a GLH biosensor chip (Bio-Rad). After the injection of IDO, the surface was blocked with 1 M ethanol-amine. The final immobilization level for IDO was approximately 18000 RU (Fig. S2†). The compound was prepared in phosphate buffered saline containing 0.005% Tween-20 (pH 7.4) and injected at 20 mL min⁻¹ for 180 s at concentrations of 1–10 μ M (1:2 dilutions). Following compound injection, the chip surface was regenerated with 30 s pulses of running buffer. The data collected were reference-subtracted using ProteOn ManagerTM 2.0.

4.10. Molecular dynamics (MD) simulations

The crystal structure of IDO (PDB access code: 5EK3) was selected from the RCSB PDB database.^{22,42,43} The crystal structure was used to explore the binding modes between the inhibitors and IDO preliminarily. Schrödinger software 2015.02 (Schrödinger Inc., New York, USA) was utilized to process the protein and the active compounds as ligands. Glide was used to dock the ligand structures into IDO with some adjustment of the coordination bond distance.⁴⁴ After docking a ligand into the IDO crystal structure, MD simulations were carried out to explore the binding modes. The protein-ligand complexes were prepared using the protocol. The MD simulations were carried out using the AMBER16 molecular simulation package. The combined free energy (ΔG) between the ligand and IDO was calculated.⁴⁵ A detailed preparation procedure for the inhibitors and IDO protein as well as details of the MD simulations can be found in the ESI.†

Conflicts of interest

The authors declare no competing interests.

Acknowledgements

This work was supported by the National Science Foundation of China (81473138), the Guangdong Frontier & Key Technology Innovation Program (2015B010109004), Guangdong NSF (2016A030310228), the Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation (pdjh2017a0005), and the Special Program for Applied Research on Super Computation of the NSFC-Guangdong Joint Fund (the second phase) under Grant No. U1501501.

References

- 1 A. W. Yeung, A. C. Terentis, N. J. King and S. R. Thomas, *Clin. Sci.*, 2015, **129**, 601–672.
- 2 J. M. Gostner, K. Becker, F. Uberall and D. Fuchs, *Expert Opin. Ther. Targets*, 2015, **19**, 605–615.

- 3 L. Zhai, S. Spranger, D. C. Binder, G. Gritsina, K. L. Lauing, F. J. Giles and D. A. Wainwright, *Clin. Cancer Res.*, 2015, **21**, 5427–5433.
- 4 P. Puccetti and U. Grohmann, *Nat. Rev. Immunol.*, 2007, **7**, 817–823.
- 5 H. Barth and S. Raghuraman, *Crit. Rev. Microbiol.*, 2014, **40**, 360–368.
- 6 S. Kim, B. J. Miller, M. E. Stefanek and A. H. Miller, *Cancer*, 2015, **121**, 2129–2136.
- 7 D. H. Munn and A. L. Mellor, *Trends Immunol.*, 2013, **34**, 137–143.
- 8 A. Curti, S. Trabanelli, V. Salvestrini, M. Baccarani and R. M. Lemoli, *Blood*, 2009, **113**, 2394–2401.
- 9 D. H. Munn and A. L. Mellor, *J. Clin. Invest.*, 2007, **117**, 1147–1154.
- 10 D. H. Munn and A. L. Mellor, *Trends Immunol.*, 2016, **37**, 193–207.
- 11 G. Z. Reus, K. Jansen, S. Titus, A. F. Carvalho, V. Gabbay and J. Quevedo, *J. Psychiatr. Res.*, 2015, **68**, 316–328.
- 12 H. Fujigaki, Y. Yamamoto and K. Saito, *Neuropharmacology*, 2017, **112**, 264–274.
- 13 M. Mailankot, M. M. Staniszevska, H. Butler, M. H. Caprara, S. Howell, B. Wang, C. Doller, L. W. Reneker and R. H. Nagaraj, *Lab. Invest.*, 2009, **89**, 498–512.
- 14 A. M. Myint and Y. K. Kim, *Prog. Neuro-Psychopharmacol. Biol. Psychiatry*, 2014, **48**, 304–313.
- 15 L. C. Souza, C. R. Jesse, M. S. Antunes, J. R. Ruff, D. de Oliveira Espinosa, N. S. Gomes, F. Donato, R. Giacomeli and S. P. Boeira, *Brain, Behav., Immun.*, 2016, **56**, 363–377.
- 16 S. Yang, X. Li, F. Hu, Y. Li, Y. Yang, J. Yan, C. Kuang and Q. Yang, *J. Med. Chem.*, 2013, **56**, 8321–8331.
- 17 C. J. Yu, M. F. Zheng, C. X. Kuang, W. D. Huang and Q. Yang, *J. Alzheimer's Dis.*, 2010, **22**, 257–266.
- 18 E. W. Yue, B. Douty, B. Wayland, M. Bower, X. Liu, L. Leffet, Q. Wang, K. J. Bowman, M. J. Hansbury, C. Liu, M. Wei, Y. Li, R. Wynn, T. C. Burn, H. K. Koblish, J. S. Fridman, B. Metcalf, P. A. Scherle and A. P. Combs, *J. Med. Chem.*, 2009, **52**, 7364–7367.
- 19 E. Vacchelli, F. Aranda, A. Eggermont, C. Sautes-Fridman, E. Tartour, E. P. Kennedy, M. Platten, L. Zitvogel, G. Kroemer and L. Galluzzi, *OncoImmunology*, 2014, **3**, e957994.
- 20 X. Liu, N. Shin, H. K. Koblish, G. Yang, Q. Wang, K. Wang, L. Leffet, M. J. Hansbury, B. Thomas, M. Rupa, P. Waeltz, K. J. Bowman, P. Polam, R. B. Sparks, E. W. Yue, Y. Li, R. Wynn, J. S. Fridman, T. C. Burn, A. P. Combs, R. C. Newton and P. A. Scherle, *Blood*, 2010, **115**, 3520–3530.
- 21 S. Serra, L. Moineaux, C. Vancraeynest, B. Masereel, J. Wouters, L. Pochet and R. Frederick, *Eur. J. Med. Chem.*, 2014, **82**, 96–105.
- 22 Y. H. Peng, S. H. Ueng, C. T. Tseng, M. S. Hung, J. S. Song, J. S. Wu, F. Y. Liao, Y. S. Fan, M. H. Wu, W. C. Hsiao, C. C. Hsueh, S. Y. Lin, C. Y. Cheng, C. H. Tu, L. C. Lee, M. F. Cheng, K. S. Shia, C. Shih and S. Y. Wu, *J. Med. Chem.*, 2016, **59**, 282–293.
- 23 U. F. Rohrig, S. R. Majjigapu, M. Chambon, S. Bron, L. Pilotte, D. Colau, B. J. Van den Eynde, G. Turcatti, P. Vogel, V. Zoete and O. Michielin, *Eur. J. Med. Chem.*, 2014, **84**, 284–301.
- 24 J. R. Smith, K. J. Evans, A. Wright, R. D. Willows, J. F. Jamie and R. Griffith, *Bioorg. Med. Chem.*, 2012, **20**, 1354–1363.
- 25 A. Gaulton, L. J. Bellis, A. P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani and J. P. Overington, *Nucleic Acids Res.*, 2012, **40**, D1100–D1107.
- 26 C. E. Blunt, C. Torcuk, Y. Liu, W. Lewis, D. Siegel, D. Ross and C. J. Moody, *Angew. Chem., Int. Ed.*, 2015, **54**, 8740–8745.
- 27 A. B. Dounay, J. B. Tuttle and P. R. Verhoest, *J. Med. Chem.*, 2015, **58**, 8762–8782.
- 28 S. Fallarini, A. Massarotti, A. Gesu, S. Giovarruscio, G. C. Zabetta, R. Bergo, B. Giannelli, A. Brunco, G. Lombardi, G. Sorba and T. Pirali, *MedChemComm*, 2016, **7**, 409–419.
- 29 W. P. Malachowski, M. Winters, J. B. DuHadaway, A. Lewis-Ballester, S. Badir, J. Wai, M. Rahman, E. Sheikh, J. M. LaLonde, S. R. Yeh, G. C. Prendergast and A. J. Muller, *Eur. J. Med. Chem.*, 2016, **108**, 564–576.
- 30 U. F. Rohrig, S. R. Majjigapu, P. Vogel, V. Zoete and O. Michielin, *J. Med. Chem.*, 2015, **58**, 9421–9437.
- 31 J. Xu, *J. Med. Chem.*, 2002, **45**, 5311–5320.
- 32 L. Wang, X. Le, L. Li, Y. Ju, Z. Lin, Q. Gu and J. Xu, *J. Chem. Inf. Model.*, 2014, **54**, 3186–3197.
- 33 Z. Liu, M. Zheng, X. Yan, Q. Gu, J. Gasteiger, J. Tijhuis, P. Maas, J. Li and J. Xu, *J. Comput.-Aided Mol. Des.*, 2014, **28**, 941–950.
- 34 S. Y. Yang, *Drug Discovery Today*, 2010, **15**, 444–450.
- 35 Y. Li, L. Wang, Z. Liu, C. Li, J. Xu, Q. Gu and J. Xu, *Mol. Biosyst.*, 2015, **11**, 1241–1250.
- 36 J. Fang, Y. Li, R. Liu, X. Pang, C. Li, R. Yang, Y. He, W. Lian, A. L. Liu and G. H. Du, *J. Chem. Inf. Model.*, 2015, **55**, 149–164.
- 37 T. K. Littlejohn, O. Takikawa, D. Skylas, J. F. Jamie, M. J. Walker and R. J. Truscott, *Protein Expression Purif.*, 2000, **19**, 22–29.
- 38 A. Matin, I. M. Streete, I. M. Jamie, R. J. Truscott and J. F. Jamie, *Anal. Biochem.*, 2006, **349**, 96–102.
- 39 C. Vecchi, G. Montosi and A. Pietrangelo, *Hepatology*, 2010, **51**, 654–659.
- 40 S. Y. Lin, T. K. Yeh, C. C. Kuo, J. S. Song, M. F. Cheng, F. Y. Liao, M. W. Chao, H. L. Huang, Y. L. Chen, C. Y. Yang, M. H. Wu, C. L. Hsieh, W. Hsiao, Y. H. Peng, J. S. Wu, L. M. Lin, M. Sun, Y. S. Chao, C. Shih, S. Y. Wu, S. L. Pan, M. S. Hung and S. H. Ueng, *J. Med. Chem.*, 2016, **59**, 419–430.
- 41 P. R. Twentyman and M. Luscombe, *Br. J. Cancer*, 1987, **56**, 279–285.
- 42 H. Sugimoto, S. I. Oda, T. Otsuki, T. Hino, T. Yoshida and Y. Shiro, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 2611–2616.
- 43 S. Tojo, T. Kohno, T. Tanaka, S. Kamioka, Y. Ota, T. Ishii, K. Kamimoto, S. Asano and Y. Isobe, *ACS Med. Chem. Lett.*, 2014, **5**, 1119–1123.
- 44 *Glide, version 5.6*, Schrödinger, LLC, New York, 2009.

- 45 Z. Liu, J. Zhou, R. Wu and J. Xu, *J. Chem. Theory Comput.*, 2014, **10**, 5057–5067.
- 46 L. Zhou, Z. Zuo and M. S. Chow, *J. Clin. Pharmacol.*, 2005, **45**, 1345–1359.
- 47 S. Liu, H. Xin, Y. Nie, A. Trivett, J. J. Oppenheim and D. Yang, *J. Immunol.*, 2016, **196**, 142.20.
- 48 Z. Xu, H. Jiang, Y. Zhu, H. Wang, J. Jiang, L. Chen, W. Xu, T. Hu and C. H. Cho, *Chem.-Biol. Interact.*, 2017, **273**, 48–55.