



The discovery of novel small molecule allosteric activators of aldehyde dehydrogenase 2



Wei Tian ^{a, b, 1}, Jiapeng Guo ^{a, 1}, Qingsen Zhang ^{a, 1}, Shaoyu Fang ^a, Ruolan Zhou ^a, Jian Hu ^a, Mingping Wang ^a, Yuefan Zhang ^c, Jin-Min Guo ^d, Zhuo Chen ^e, Ju Zhu ^{a, **}, Canhui Zheng ^{a, *}

^a School of Pharmacy, Second Military Medical University, Shanghai, 200433, China

^b General Hospital Of Central Theater Command, Wuhan, Hubei, 430070, China

^c School of Medicine, Shanghai University, Shanghai, 20444, China

^d 960 Hospital of the Joint Logistics Support Force, Jinan, Shandong, 250031, China

^e School of Pharmacy, East China University of Science and Technology, Shanghai, 200237, China

ARTICLE INFO

Article history:

Received 21 September 2020

Received in revised form

1 December 2020

Accepted 17 December 2020

Available online 28 December 2020

Keywords:

Aldehyde dehydrogenase 2

Allosteric activators

Inactive Asian variant

Virtual screening

Structural optimization

Ischemic stroke

ABSTRACT

Aldehyde dehydrogenase 2 (ALDH2) plays important role in ethanol metabolism, and also serves as an important shield from the damage occurring under oxidative stress. A special inactive variant was found carried by 35–45% of East Asians. The variant carriers have recently been found at the higher risk for the diseases related to the damage occurring under oxidative stress, such as cardiovascular and cerebrovascular diseases. As a result, ALDH2 activators may potentially serve as a new class of therapeutics. Herein, *N*-benzylanilines were found as novel allosteric activators of ALDH2 by computational virtual screening using ligand-based and structure-based screening parallel screening strategy. Then a structural optimization was performed and has led to the discovery of the compound **C6**. It has good activity *in vitro* and *in vivo*, which could reduce infarct size by ~70% in ischemic stroke rat models. This study provided good lead compounds for the further development of ALDH2 activators.

© 2020 Elsevier Masson SAS. All rights reserved.

1. Introduction

Aldehyde dehydrogenase (ALDH), one of the most important superfamilies of phase I oxidizing enzymes besides P-450 superfamily, is responsible for the detoxification of biogenic and xenogenic aldehydes [1,2]. ALDH2, one of 19 ALDHs, is well known for its role in detoxifying acetaldehyde generated from ethanol metabolism. ALDH2 also plays a key role in oxidizing endogenous aldehydic products that arise from lipid peroxidation under oxidative stress, such as 4-hydroxy-2-nonenal (4-HNE) and malondialdehyde (MDA) [1,3]. Therefore, it serves as an important shield from the damage occurring under oxidative stress.

A special variant ALDH2*2 with reduced enzymatic activity was found originated from Chinese thousands years ago. It is carried by

the highest percentage of the population in Southeast China (up to 65% in Changting County of Fujian province), and 35–45% of East Asians, while lower than 5% of Caucasians [4,5]. This variant has recently more and more been recognized as having a major impact on human health and disease, besides its effect on limiting alcohol consumption. The variant carrier drinkers are at the higher risk for several cancers [3,6,7]. In addition, even the variant carrier non-drinkers are also at the higher risk for the diseases related to the damage occurring under oxidative stress, such as cardiovascular and cerebrovascular diseases, diabetes, neurodegenerative disease, Fanconi anemia, pain, and osteoporosis [1,3,8]. In particular, that an increased incidence of cardiovascular and cerebrovascular diseases (myocardial ischemia [9–11] or cerebral ischemia [12–15]) is associated with the variant was discovered. As a result, ALDH2 activators may potentially serve as a new class of therapeutics, and they can also serve as important molecular probes to further explore the function of ALDH2.

ALDH2 is a tetramer of four identical subunits, and each is composed of three main domains: the catalytic domain, the

* Corresponding author.

** Corresponding author.

E-mail address: canhuizheng@smmu.edu.cn (C. Zheng).

¹ These authors contributed equally.

coenzyme or NAD⁺-binding domain (Fig. 1), and the oligomerization domain. The ALDH2*2 mutation occurs at amino acid residue 487 locating within the oligomerization domain, where a basic amino acid lysine substituted for an acidic amino acid glutamic acid (E487K). This single amino acid substitution disorders a large region located at the dimer interface, causing an allosteric disruption of the catalytic and coenzyme-binding sites. This structural instability of the ALDH2*2 results in the reduced enzymatic activity compare with wide-type ALDH2 [3,16–18]. The only well studied family of ALDH2 activators, represented by Alda-1, was discovered using a high-throughput screening [11]. Cocrystal structures of activator Alda-1 with wild-type ALDH2 or ALDH2*2 (Fig. 1) demonstrated that Alda-1 binds at the entrance of the catalytic domain, and the binding of Alda-1 restores the disordered region nearly to the native wildtype state even though Alda-1 has no direct contact with these residues. Alda-1 is therefore an activator which exerts its allosteric effect to restore the structural defect of a catalytically impaired enzyme [1,16]. It could increase the activity of ALDH2*2 by ~ 10 folds, and ALDH2 by ~ 2 folds [11,15]. Previously, structural modification of Alda-1 with *N*-benzylbenzamide scaffold was conducted by us. Several new derivatives with improved water-solubility were obtained, while no ALDH2 activator with better activity than Alda-1 was found [19].

Nowadays, computational virtual screening is becoming a powerful tool for the identification of lead compound in drug discovery [20–23]. However, only a few successful examples of the virtual screening of the allosteric activators have been reported. The reason maybe is the binding of allosteric activators often results in a conformational change of the target protein, which reduces the accuracy of molecular docking and then increases the difficulty of structure-based virtual screening. In this study, taking advantage of complementary ligand-based and structure-based virtual screening techniques and overcoming their limitations, a parallel screening strategy was used. It identified *N*-benzylanilines as a novel family of allosteric activators of ALDH2. Then a structural optimization was performed and has led to the discovery of the lead compound **C6** with good activity *in vitro* and *in vivo*.

2. Results and discussion

2.1. First round of virtual screening: pharmacophore-based and structure-based parallel screening

3D Pharmacophores are widely used for virtual screening of databases in order to identify new lead series [24]. The Receptor–Ligand Pharmacophore Generation (RLPG) Protocol generates pharmacophore models directly from the receptor–ligand interactions as revealed in the 3-D structures, which are recommend for profiling targets for which only a few ligands but a 3-D structure is available [25,26]. RLPG was used to generate the pharmacophore model of the activators of ALDH2. The model identified mainly five functional features, including one ring aromatic region, three hydrophobic regions, hydrogen-bond donor, and several exclusion regions (Fig. 1). The ring aromatic region feature overlaps the phenyl group of Alda-1 binding deep into the ALDH2. This feature is consistent with the Alda-1-ALDH2 interaction mode: one phenyl group shows π - π stacking interactions with Phe459, Phe170 and Phe296; The three hydrophobic regions overlap with the methylene in dioxolane group, the other phenyl group and a chlorine group of Alda-1, respectively, which show hydrophobic interactions with the active site; The hydrogen-bond donor feature overlap with the nitrogen in the amide group that have been reported to interact with the backbone of Asp457. To validate the model, a pharmacophore-based virtual screening was undertaken on a test database, which contains Alda-1 and 200 compounds random selected from SPECS database. Unfortunately, 15 compounds got higher score than Alda-1, which indicated the low efficiency of this screening. After analysis, because of the catalytic site is adjacent to the binding site of ALDH2 activators most of the false positive compounds bound deep into it, which should be the binding site of inhibitors. Therefore, three exclusion region features (Fig. 1) were added manually to improve the effectiveness of virtual screening. Using the new pharmacophore model, Alda-1 successfully got the highest score in the test screening.

Additionally, Glide was selected as the docking method for virtual screening by comparing four docking programs [27]. This method was further validated by the test screening described above, which

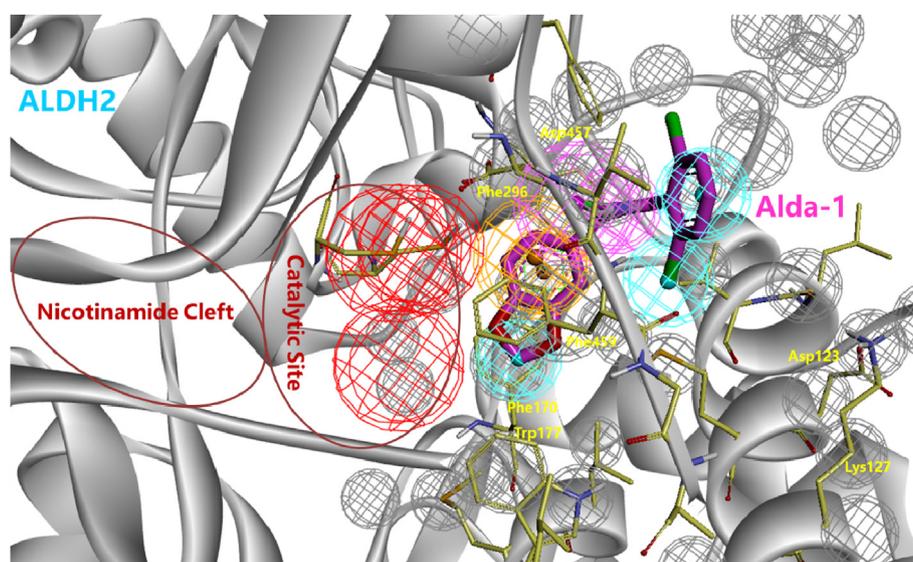


Fig. 1. Binding modes of Alda-1 with ALDH2 and receptor-ligand based pharmacophore model of ALDH2 (crystal structure PDB ID: 3INL). Alda-1 is showed as purple sticks. The catalytic domain and the NAD⁺-binding domain are showed inside the red circles. The model included one ring aromatic region (orange sphere), three hydrophobic regions (cyan spheres), hydrogen-bond donor (magenta sphere), several exclusion regions (gray spheres) and three exclusion regions added manually (red spheres). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

gave Alda-1 second highest score, which indicated that it was also a suitable method for structure-based virtual screening. Then in the first round, a pharmacophore-based and structure-based parallel screening was carried out to search the SPECS database [22], the schematic of which is presented in Fig. 2A. After extra precision docking, visual inspection and with the consideration of the chemical diversity, 55 candidates were selected and purchased finally from top ranked compounds (Table S1).

Considered that the ALDH2 activators were reported to have the activation activity both on ALDH2 and ALDH2*2,¹¹ and only ALDH2 protein is commercially available, the activities of the title compounds were evaluated in a fluorescent ALDH2 enzyme assay [11]. The maximum fold activation was taken as the primary indicator, and the activities were corrected by the control Alda-1. The maximum fold activation of Alda-1 was 1.6–1.8 in our different tests, which was assigned as 100%. All purchased compounds were evaluated, and 2 hits showed detectable activities (Table 1). Because compound **A1** has an ester group inside its scaffold which is unstable *in vivo*, compound **A2** with *N*-benzylaniline scaffold, which showed 8.8% maximum fold activation relative to Alda-1, was selected to further studies.

2.2. Second round of virtual screening: hit-based substructure search

Considered that compounds in many commercial chemical databases are inherently enriched with analogues, a structure-activity relationship (SAR) of the hits identified through virtual screening could easily be established and better hits could be found through subsequent extraction of analogues in the same database [20]. Then the second round hit-based substructure search, another ligand-based screening, was undertaken to search the Specs database (showed in Fig. 2B). Because 3,4-dichlorophenyl group of **A2** was predicted to bind to ALDH2 in the same site with the 1,3-benzodioxole group of Alda-1, which has better binding capacity, *N*-(benzo[*d*][1,3]dioxol-5-ylmethyl)aniline (showed in Fig. 2 as **A2** scaffold) was used as the scaffold of the following substructure search. 6 Analogues of **A2** were selected and purchased for biological evaluation finally. The activation activities of all compounds were evaluated and displayed in Table 2.

Replacement of the 3,4-dichlorophenyl group of **A2** with 1,3-

benzodioxole group resulted in compound **B1**, which had slightly better activation activity than **A2**. Replacement of morpholine or carboxylic acid group with chlorine group got compounds **B2** or **B3** respectively, and replacement or removing of these two groups simultaneously got compounds **B4** or **B5**. These four compounds could not show detectable activities, indicating that morpholine and carboxylic acid groups were important for their activation activity. In addition, introduction of ethyl ester group on the carboxylic acid group of compound **B1** obtained compound **B6**, which showed much better activation activity than **B1** and **A2**, indicating that hydrophobic substitutes on the carboxylic acid group were desirable for achieving a better activity. Compound **B6** displayed 63.7% maximum fold activation corrected by that of Alda-1, and it could be a good starting point for further optimization studies.

2.3. Structural optimization

The docking mode of **B6** in complex with ALDH2 (Fig. 3B) was obtained using crystal structure 3INJ as template. One phenyl group showed π - π stacking interactions and the nitrogen in the benzylamino group formed hydrogen-bond with ALDH2, similar with the same groups of Alda-1 (Fig. 3A). However, the other groups of **B6** displayed quite different binding mode to ALDH2 with Alda-1: the ester group showed hydrogen-bond with the backbone of Phe459 as a hydrogen-bond acceptor. The ethyl group in the ester group formed hydrophobic interactions with a narrow hydrophobic pocket on the surface of ALDH2. The morpholine group bound to a polar pocket on the surface of ALDH2. These two sites haven't been utilized by Alda-1 for binding. This docking mode was consistent with the preliminary SAR of this type of compound from the second round screening, and could guide the further structural optimization of **B6** (Table 3).

Considered that the ester group is unstable *in vivo*, the replacement of it of **B6** with amide group was performed. In addition, the 1,3-benzodioxole group of Alda-1 could be replaced by 3-fluoro-4-methoxyphenyl group with more stability [28]. Considered that the same group of **B6** was predicted to bind to ALDH2 in the same site, the replacement was also carried out to generate compound **C1**. It had comparable activation activity with **B6**, and was used as the new template compound for subsequent modifications.

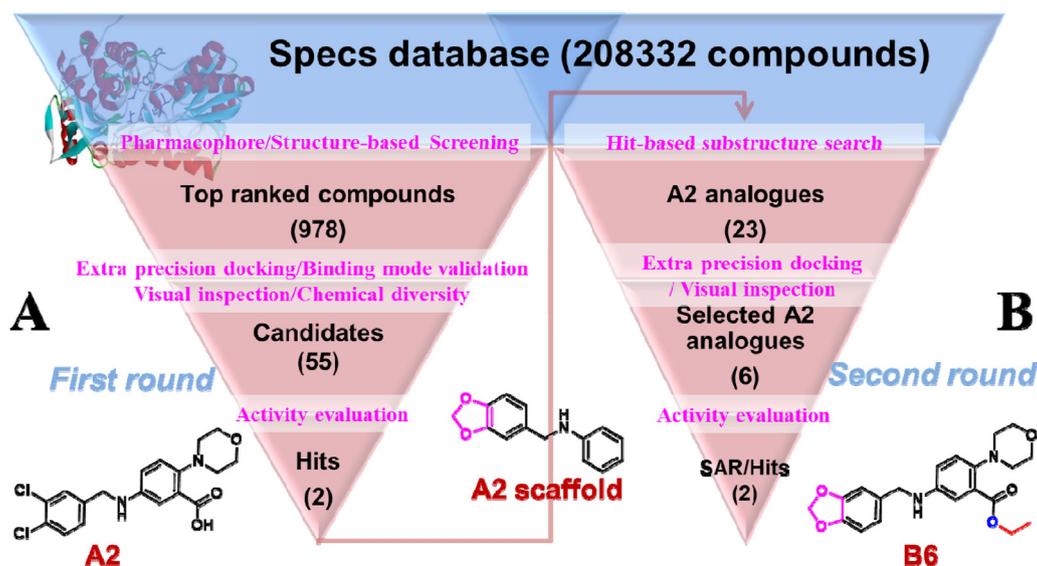


Fig. 2. The strategy of two rounds of virtual screening: pharmacophore-based and structure-based parallel screening (A), and hit-based substructure search (B).

Table 1
Structures and activation activities of the hits obtained from the first round of virtual screening.

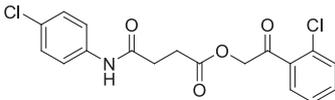
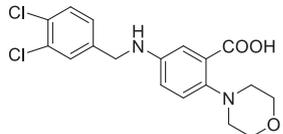
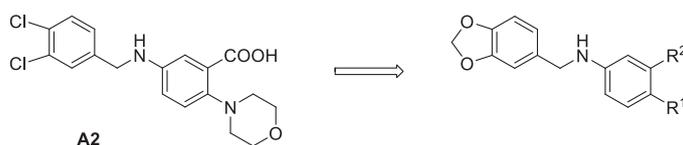
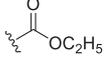
No.	Structures	Maximum fold activation corrected by that of Alda-1 (%)
A1		6 ± 1
A2		9 ± 2

Table 2
Structures and activation activities of the hits obtained from the second round of virtual screening.



No.	Structures		Maximum fold activation corrected by that of Alda-1 (%)
	R [1]	R [2]	
B1			23 ± 1
B2			N. A.
B3			N. A.
B4			N. A.
B5		/	N. A.
B6			64 ± 2

N. A.: No activity.

Firstly, we performed modifications of the R¹ group on compound **C1** (Table 3). A series of compounds **C2** – **C11** containing different aliphatic R¹ groups were synthesized. Compounds **C3**, **C4**, **C5** and **C6** with substitutes containing 3–4 carbons had good activation activities. Compound **C6** showed 104.1% maximum fold activation corrected by that of Alda-1. These substitutes may be the groups which fitted with the R¹ binding pocket of ALDH2 very well (Fig. 3C). Then, the fluorine or hydroxyl group was introduced into R¹ of **C5** and **C6** to get compounds **C12**, **C13** and **C14**, which also had good activation activities. The compound **C14** had the best activity, with 119.9% maximum fold activation corrected by that of Alda-1. The hydroxyl group of **C14** was predicted to stretch to the opening of the active site, and might form hydrogen bonds potentially with the side chain of Lys127 considered its flexibility (Fig. 3D).

We next employed compound **C6** as the template compound for further modifications of R² group (Table 3). Removing of the

morpholine group or the nitrogen atom in it got compounds **C15** and **C20**, which displayed reduced activation activity than **C6**. In addition, replacement of the morpholine group by phenyl, or the oxygen atom by methylene with or without fluorine substituted resulted in compounds **C21**, **C16** – **C18**. These compounds showed very low or no detectable activity. However, compound **C19** generated from moving the oxygen atom out of the morpholine ring, displayed slightly weaker activity than **C6**. Above results indicated that the morpholine group, especially the distal polar atom, was important for the activity of this type of compound.

Finally, we introduced methyl or acetyl as R³ group on the nitrogen in the benzylamino group of compounds **C6**, **C13** and **C14** to get compounds **C22**, **C23** and **C24** respectively, which displayed no detectable activity. It was consistent with the predicted binding model, in which the nitrogen atom of this type of compound formed important hydrogen-bonds with ALDH2. In addition, we

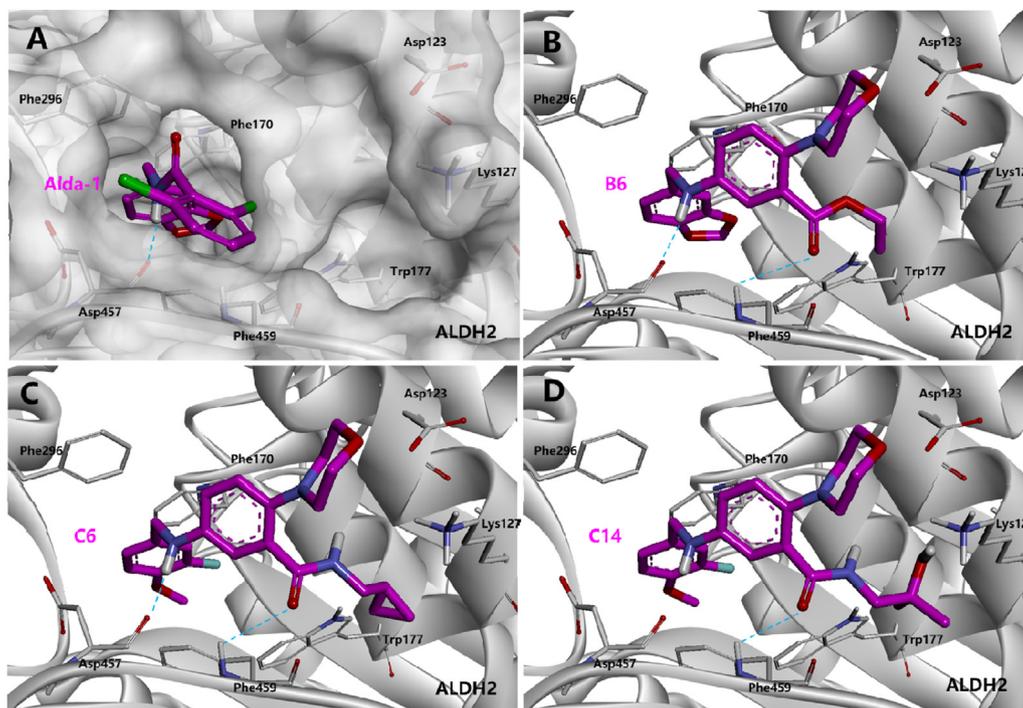


Fig. 3. Binding modes of Alda-1 (A) and docking mode of **B6** (B), **C6** (C) and **C14** (D) in complex with ALDH2 (crystal structure PDB ID: 3INJ). Ligands are shown as purple sticks. ALDH2 is shown in solvent-accessible surface or gray ribbons representation, and important interacting residues are shown in gray sticks. Hydrogen bonds are shown as cyan dotted lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

replaced the methylene in the benzylamino group of compound **C6** with carbonyl group to yield compound **C25**. It also showed no detectable activity, which indicated that the *N*-benzylaniline scaffold of this type of compound was essential for their activity.

2.4. Further biological evaluation of **C6**

Representative ALDH2 activator **C6** showed good dose-response relationship in the ALDH2 activation assay for different concentrations (Fig. 4A). The EC_{50} value could be determined from the curve as the second indicator of the activation activities. Compound **C6** had the EC_{50} value of $17.9 \pm 1.0 \mu\text{M}$, which slight weaker than that of Alda-1 ($4.3 \pm 1.0 \mu\text{M}$). Furthermore, to preliminarily evaluate its effect *in vivo*, middle cerebral artery occlusion (MCAO) Sprague-Dawley (SD) rat model was used, which could simulate ischemic stroke-like events [15]. A potent ALDH2 activator **C6** was intracerebroventricular (i.c.v.) injected and the infarct size of the brain was measured by the TTC staining assay. The injection of **C6** ($64 \mu\text{g}$ in $5 \mu\text{l}$) could reduce infarct size significantly by $\sim 70\%$ (Fig. 4B and C), which achieved comparable good effect with Alda-1 in the same molar dosage ($50 \mu\text{g}$ in $5 \mu\text{l}$). In conclusion, this type of compound showed good activity *in vitro* and *in vivo*, therefore they could be the valuable lead compound of ALDH2 activator for further research and development.

2.5. Chemistry

The compounds in Table 3 were synthesized using the method outlined in Scheme 1. Firstly, 2-fluoro-5-nitrobenzoic acid (**1a**) or 3-nitrobenzoic acid (**1b**) was esterified by SOCl_2 in EtOH to get compound **2a** or **2b**. The fluorine group of **2a** was displaced by different amines by heating with K_2CO_3 in DMSO to produce compounds **I**. Then the nitro group of compounds **2b** or **I** were reduced by heating with $\text{Fe}/\text{NH}_4\text{Cl}$ in EtOH/ H_2O to form compounds **II**. Intermediates **III**

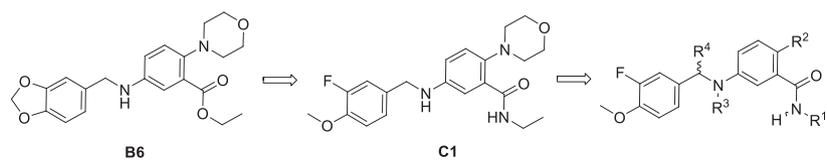
were constructed by the reductive amination reaction with 3-fluoro-4-methoxybenzaldehyde. Then they were hydrolyzed using NaOH in MeOH/THF/ H_2O to obtain the key intermediates **IV**. Finally, title compounds **C1–C19** were got by condensation with different amines under the condition of EDCI/DMAP/DCM. Another title compound **C22** was obtained by reductive amination reaction with formaldehyde aqueous solution from compound **C6**. Title compounds **C23** and **C24** were obtained by acylation reaction with acetic anhydride from compounds **C13** and **C14** separately.

2-Bromine-5-nitrobenzoic acid (**3**) was condensation with cyclopropane methyl amine to produce compound **4**. Then the bromine group of **4** was displaced by tetrahydropyran group by an onepot Suzuki-Miyaura-hydrogenation reaction to get compound **5** [29]. It could also be displaced by phenyl group by Suzuki-Miyaura coupling reaction and a following reduction reaction to get compound **7**. Finally, title compounds **C20–C21** were constructed by the reductive amination reaction with 3-fluoro-4-methoxybenzaldehyde separately. In addition, compound **1a** was reacted with morpholine to produce compound **8**. Then, compound **9** were got by condensation with cyclopropane methyl amine under the condition of HATU/DIPEA/DMF. It was reduced to form compound **10**. Finally, title compound **C25** was got by condensation with 3-fluoro-4-methoxybenzoic acid.

3. Conclusion

In this study, *N*-benzylanilines were found as novel allosteric activators of ALDH2 by two rounds of virtual screening, pharmacophore-based and structure-based parallel screening, and hit-based substructure search. Then a structural optimization was performed and has led to the discovery of compound **C6** with good activity *in vitro* and *in vivo*. It could reduce infarct size by $\sim 70\%$ in ischemic cerebral injury rat models. This study provided good lead compounds for the further development of the ALDH2 activator,

Table 3
Structures and activation activities of the compounds generated from the structural optimization..



No.	Structures				Maximum fold activation corrected by that of Alda-1 (%)
	R [1]	R [2]	R [3]	R [4]	
C1			-H	-H	63 ± 1
C2			-H	-H	52 ± 1
C3			-H	-H	94 ± 3
C4			-H	-H	97 ± 6
C5			-H	-H	92 ± 0.0
C6			-H	-H	104 ± 2
C7			-H	-H	59 ± 2
C8			-H	-H	41 ± 10
C9			-H	-H	51 ± 2
C10			-H	-H	80 ± 5
C11			-H	-H	66 ± 12
C12			-H	-H	101 ± 6
C13			-H	-H	97 ± 6
C14			-H	-H	120 ± 8
C15		-H	-H	-H	56 ± 5
C16			-H	-H	N. A.
C17			-H	-H	N. A.
C18			-H	-H	12 ± 1

Table 3 (continued)

No.	Structures				Maximum fold activation corrected by that of Alda-1 (%)
	R [1]	R [2]	R [3]	R [4]	
C19			-H	-H	80 ± 4
C20			-H	-H	53 ± 3
C21			-H	-H	3 ± 7
C22			-CH ₃	-H	N. A.
C23				-H	N. A.
C24				-H	N. A.
C25			-H		N. A.

N. A.: No activity.

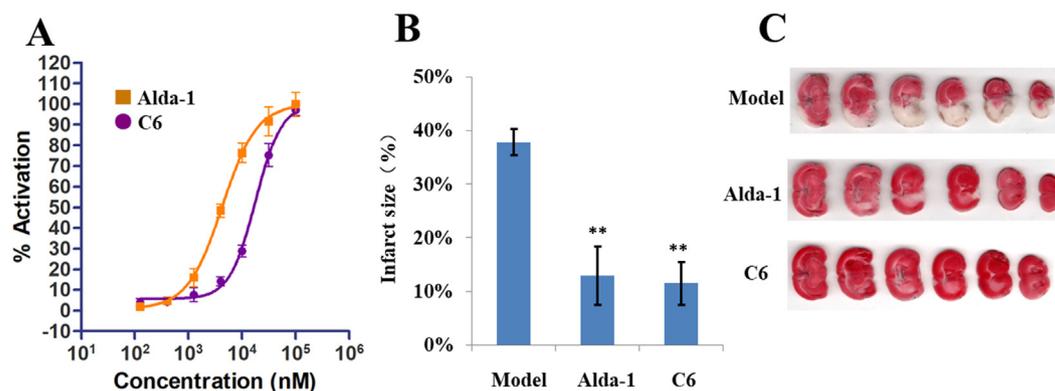


Fig. 4. (A) Dose-response curve of assay for ALDH2 activation for **C6** and Alda-1. The activities were showed as the maximum fold activation corrected by the activities of Alda-1; (B) The protective effect of **C6** and Alda-1 with same molar dosage against ischemic stroke in MCAO male rats. The infarct size in MCAO rats are shown as mean ± SD, **P < 0.01, n = 6; (C) Representative 1% TTC staining of 6 corresponding coronal brain sections of rats with vehicle, **C6** and Alda-1 treatment. The ischemic infarct region is white.

further extensive SAR study was ongoing. The potent ALDH2 activators would serve as a new class of therapeutics for several diseases related to the damage occurring under oxidative stress, and they could also serve as important molecular probes to further explore the function of ALDH2.

4. Experimental section

4.1. Molecular modeling

4.1.1. Pharmacophore model construction

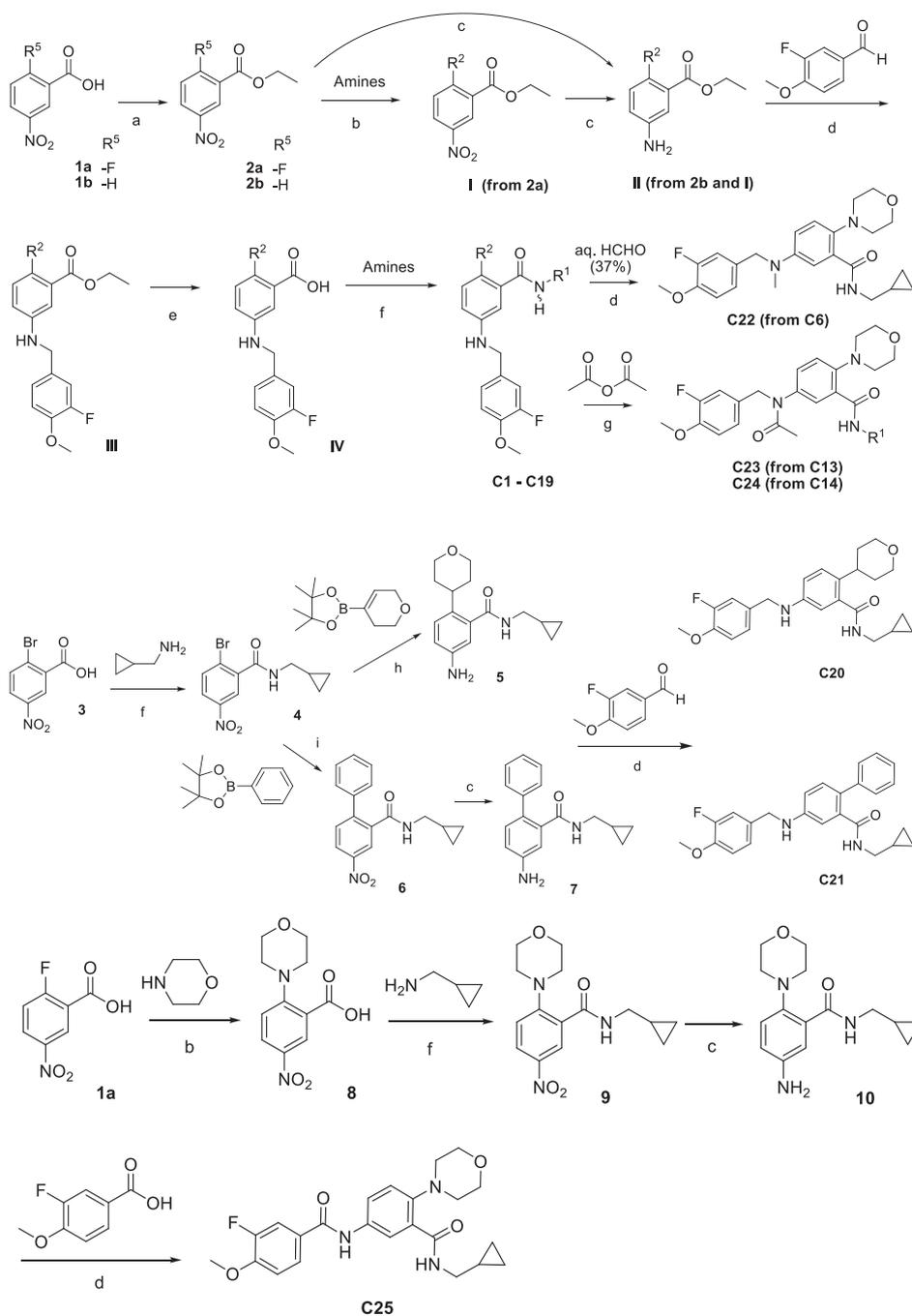
The X-ray crystal structure of ALDH2-Alda-1 complex was obtained from the Brookhaven Protein Databank entry 3INL. All crystallographic water molecules were removed from the coordinate set. The pharmacophore model of ALDH2 activators was constructed on Discovery Studio software package [30]. Most of the functional features in pharmacophore model were generated by the Receptor-Ligand Pharmacophore Generation Protocol. The exclusion region features were generated by the Exclusion Constraint Protocol and clustered by the Cluster Current Features Protocol or

manual mode. The models were validated by a pharmacophore-based virtual screening on a test database, which contains Alda-1 and 200 compounds random selected from SPECS database.

4.1.2. Molecular docking

The ligand Alda-1 extracted from the crystal structure complex 3INL was docked back into ALDH2 by Libdock [30], Gold [31], MVD [32] and Glide [27]. The highest scoring conformations in the docking results were compared with the original conformation within the crystal structure complex to calculate the RMSD values, which showed that Glide and MVD offered the best results (0.3933 and 0.4191 respectively). Then, these two methods were further validated by the test screening described above, which gave Alda-1 second and 28th highest score respectively. The result indicated that Glide was the better method for structure-based virtual screening in this study.

Docking runs of Libdock were performed using the Receptor-ligand interactions Protocol on Discovery Studio. The Fast conformation method and high quality docking preferences, and best conformation method were used, and other parameters were



Scheme 1. Synthesis of Title Compounds. Reagents and conditions: (a) $\text{SOCl}_2/\text{EtOH}$, r. t., 98%; (b) K_2CO_3 , DMSO, 80°C , 95%; (c) $\text{Fe}/\text{NH}_4\text{Cl}$, $\text{EtOH}/\text{H}_2\text{O}$, 80°C , ~ 80%; (d) $\text{NaBH}(\text{OAc})_3$, DCE, 80°C to r. t., 80–85%; (e) NaOH , $\text{MeOH}/\text{THF}/\text{H}_2\text{O}$, r. t., ~ 90%; (f) EDCI/DMAP , DCM, r. t., 70–80%; (g) DMAP , DCM, $\text{C}_5\text{H}_5\text{N}$, rt, 87%; (h) PdXPhosG2 (1 mol%), Pd/C (12mol%), K_3PO_4 (3eq), Dioxane: H_2O (4:1), 80°C , 4h; then MeOH , NH_4COOH , rt to 60°C , 16h; 75%; (i) $\text{Pd}(\text{dppf})\text{Cl}_2$, K_2CO_3 , Dioxane: H_2O (3:1), 100°C , 12h, 70%.

default; For docking runs of Gold, GoldScore was used as the scoring function, most accurate genetic algorithm search options was set, and other parameters were default; For docking runs of MVD, MolDock Score and MolDock Score [Grid] were used as scoring functions, the default settings were used; Docking runs of Glide were performed by the Ligand Docking Protocol on Schrödinger. 3 \AA around Cys301, Cys302, and Cys303 were set as excluded volumes for preparation of the active cavity grid files. Standard precision and flexible ligand sampling was employed, and other parameters were default. In addition, extra precision was used to obtain more accurate docking results.

4.1.3. Virtual screening

In the first round, pharmacophore-based and structure-based parallel screening was carried out, which completed by Screen Library Protocol installed on Discovery Studio and Glide respectively. The respective top 500 compounds obtained by two methods (978 compounds totally) was subjected to extra precision docking by Glide. Then the results were visual inspected to select the candidate compounds based on their binding mode and chemical diversity. Finally, 55 candidates were selected and purchased. The second round (hit-based substructure search was completed by Align to Selected Substructure Protocol on Discovery Studio. 23 Analogues

were found, and subjected to extra precision docking by Glide. Finally, 6 compounds were selected and purchased for biological evaluation after visual inspection. No PAINS liability was found in compounds reported by analysis in the online filter (<http://zinc15.docking.org/patterns/home>).

4.2. Chemistry

¹H NMR and ¹³C NMR spectra were recorded on Bruker AC-300P or AVANCE II600 spectrometer (Bruker Biospin, Swiss), using TMS as an internal standard and CDCl₃ or DMSO-*d*₆ as solvents. Chemical shifts are given in ppm (δ), and the spectral data are consistent with the assigned structures. The mass spectra were recorded on a Micromass Qtof-Micro LC-MS instrument. Silica gel thin-layer chromatography was performed on precoated plates GF-254. Silica gel column chromatography was performed with silica gel 60 G. All compounds were routinely checked by TLC by using silica gel plates GF-254. All solvents and reagents were analytically pure and, when necessary, were purified and dried by standard protocols. All starting materials were commercially available unless otherwise indicated. Yields of purified products were not optimized. The purity of key compounds (>95%) was determined on an Agilent 1100 series liquid chromatography (LC) system (column, Diamonsil Plus 5 μ m C18-A 4.6250 \times 4.6 mm; mobile phase, CH₃CN/H₂O; UV wavelength, absorbance at 254 nm). The candidates of virtual screening (Table S1) were purchased from the commercial source (¹H NMR and LC-MS spectra are available on the website: <http://www.specs.net>).

4.2.1. General method for step a

Compound **1a** or **1b** (10.8 mmol) was added to ethanol solution (15 ml), into which thionyl chloride (1.54 g, 12.9 mmol) was slowly dropped under ice-cooling and stirred for 5 min. Subsequently, the solution was refluxed at 80 °C for 3 h. The reaction was monitored with TLC until it completed. Then the reaction solution was concentrated by rotary evaporation and separated by silica gel column (PE/EA system) to afford compound **2a** or **2b** with 98% yield.

4.2.2. General method for step b

Compound **2a** (2 g, 9.4 mmol), appropriate amines (9.4 mmol, 1 eq.) and potassium carbonate (2.6 g, 18.8 mmol) were dissolved in 15 ml of DMSO and heated under reflux at 80 °C for 2 h. The reaction was monitored with TLC until it completed. The reaction solution was poured into ice water (300 ml), extracted with EA in triplicate, concentrated by rotary evaporation and separated by silica gel column (PE/EA system) to get compound **I** with 95% yield.

4.2.3. General method for step c

Compound **2b** or **I** (9.4 mmol) was dissolved in ethanol (15 ml), which was added with ammonium chloride (1.76 g, 32.9 mmol) dissolved in 5 ml of water. The solution was heated up to 80 °C and added with iron powder (1.76 g, 31.5 mmol) in batches. After 2 h of reaction, the reaction was confirmed by TLC to have completed. After the filtration with Celite, the filtrate was concentrated by rotary evaporation and separated via the silica gel column (PE/EA system) to afford compound **II** with ~80% yield.

4.2.4. General method for step d

Compound **II** (1.0 mmol) and aldehydes (1.0 mmol, 1.1 eq.) were dissolved in dichloroethane (10 ml), stirred at 80 °C for 1 h and allowed to cool down to room temperature. The solution was added with sodium triacetoxyborohydride (338 mg, 1.6 mmol) and allowed to react overnight at room temperature under the argon atmosphere. Completion of the reaction was confirmed by TLC. The

reaction was quenched with the saturated sodium bicarbonate solution, extracted with ethyl acetate, concentrated by rotary evaporation and separated by silica gel column (PE/EA system) to obtain compound **III** with 80–85% yield.

4.2.5. General method for step e

Compound **III** (1.0 mmol) was dissolved with 10 ml of mixed solution with the ratio of MeOH: THF at 1:1. Sodium hydroxide (400 mg, 10.0 mmol) was dissolved in 8 ml of mixed solvents with the ratio of MeOH: H₂O at 5:3, poured into the reaction mixture, and reacted at room temperature for 4 h. Completion of the reaction was confirmed by TLC. The reaction solution was concentrated by rotary evaporation and separated via a silica gel column (PE/EA system) to afford the pure compound **IV** with ~90% yield.

4.2.6. General method for step f

Compound **IV** (1 mmol), appropriate amines (1.1 mmol, 1.1 eq.), EDCI (0.96 g, 5.0 mmol) and DMAP (244.3 mg, 2.0 mmol) were dissolved in anhydrous dichloromethane and allowed to react at room temperature overnight. Completion of the reaction was confirmed by TLC. The reaction solution was washed with 1 M diluted hydrochloric acid and saturated sodium bicarbonate solution successively, concentrated by rotary evaporation and purified via a silica gel column (PE/EA system) to afford the products with 70–80% yield.

4.2.7. General method for step g

Compound **C13** or **C14** (1.0 mmol), acetic anhydride (2.0 mmol, 2 eq.), and DMAP (0.50 mmol, 0.5 eq.) was added into 5 ml pyridine. The reaction mixture was stirred for 3 h at room temperature. Completion of the reaction was confirmed by TLC. The mixture was extracted with ethyl acetate, and concentrated by rotary evaporator, then purified by silica gel column (DCM/MeOH system) to provide the products with ~80% yield.

4.2.8. Method for step h

Compound **4** (50 mg, 1.0 mmol), 2-(3,6-dihydro-2H-pyran-4-yl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane (52 mg, 1.0 mmol), PdXPhosG₂ (1 mg, 0.01 mmol), Pd/C (22 mg, 0.12 mmol) and K₃PO₄ (107 mg, 3.0 mmol) was added into 4 ml of the mixed solvent of dioxane/H₂O (3:1) under argon atmosphere. The reaction mixture was heated at 80 °C for 4 h, then cooled to room temperature. Subsequently, the reaction solution was added with ammonium formate (106 mg, 10 mmol) dissolved in methol (2 ml) and reacted at room temperature for 16 h. Completion of the reaction was confirmed by TLC. The reaction was filtered with celite, concentrated by rotary evaporation and purified by silica gel column (PE/EA system) to afford white solid with 75% yield.

4.2.9. Method for step i

Compound **4** (50 mg, 1.0 mmol), 4,4,5,5-tetramethyl-2-phenyl-1,3,2-dioxaborolane (34 mg, 1.0 mmol), Pd(dppf)Cl₂ (12 mg, 0.10 mmol), and K₂CO₃ (69 mg, 3.0 mmol) was added into 4 ml of the mixed solvent of dioxane:H₂O (3:1) under Argon atmosphere. Then, the reaction mixture was heated to 100 °C for 12 h. Completion of the reaction was confirmed by TLC. The reaction was filtered with celite, concentrated by rotary evaporation and purified by silica gel column (PE/EA system) to afford white solid with 70% yield.

4.2.10. N-ethyl-5-((3-fluoro-4-methoxybenzyl)amino)-2-morpholinobenzamide (**C1**)

White solid; ¹H NMR (600 MHz, DMSO-*d*₆) δ 9.96 (t, *J* = 5.3 Hz, 1H), 7.17–7.06 (m, 5H), 6.62 (dd, *J* = 8.7, 3.0 Hz, 1H), 6.32 (t, *J* = 6.1 Hz, 1H), 4.19 (d, *J* = 6.2 Hz, 2H), 3.79 (s, 3H), 3.76–3.66 (m,

4H), 3.31–3.27 (m, 2H), 2.80–2.75 (m, 4H), 1.15 (t, $J = 7.2$ Hz, 3H); HRMS (ESI) m/z found 388.2043 ($M+H^+$), while $C_{21}H_{26}FN_3O_3$ ($M+H^+$) requires 388.2031.

4.2.11. 5-((3-fluoro-4-methoxybenzyl)amino)-N-methyl-2-morpholinobenzamide (C2)

Light yellow solid; 1H NMR (600 MHz, DMSO- d_6) δ 9.74 (d, $J = 4.7$ Hz, 1H), 7.17–7.09 (m, 4H), 7.04 (d, $J = 8.6$ Hz, 1H), 6.62 (dd, $J = 8.6, 2.9$ Hz, 1H), 6.30 (t, $J = 6.1$ Hz, 1H), 4.18 (d, $J = 6.1$ Hz, 2H), 3.79 (s, 3H), 3.73–3.68 (m, 4H), 2.81 (d, $J = 4.8$ Hz, 3H), 2.78–2.72 (m, 4H); HRMS (ESI) m/z found 374.1876 ($M+H^+$), while $C_{20}H_{24}FN_3O_3$ ($M+H^+$) requires 374.1874.

4.2.12. N-cyclopropyl-5-((3-fluoro-4-methoxybenzyl)amino)-2-morpholinobenzamide (C3)

White solid; 1H NMR (600 MHz, DMSO- d_6) δ 10.07 (d, $J = 4.6$ Hz, 1H), 7.18–7.03 (m, 5H), 6.62 (dd, $J = 8.7, 2.9$ Hz, 1H), 6.35 (s, 1H), 4.19 (s, 2H), 3.79 (s, 3H), 3.66 (dd, $J = 5.9, 3.3$ Hz, 4H), 2.88–2.80 (m, 1H), 2.75 (t, $J = 4.5$ Hz, 4H), 0.76–0.70 (m, 2H), 0.55–0.49 (m, 2H); HRMS (ESI) m/z found 400.2028 ($M+H^+$), while $C_{22}H_{26}FN_3O_3$ ($M+H^+$) requires 400.2031.

4.2.13. 5-((3-fluoro-4-methoxybenzyl)amino)-2-morpholino-N-propylbenzamide (C4)

White solid; 1H NMR (600 MHz, DMSO- d_6) δ 10.02 (t, $J = 5.6$ Hz, 1H), 7.17 (d, $J = 3.0$ Hz, 1H), 7.16–7.06 (m, 4H), 6.62 (dd, $J = 8.7, 3.0$ Hz, 1H), 6.33 (t, $J = 6.2$ Hz, 1H), 4.19 (d, $J = 6.1$ Hz, 2H), 3.79 (s, 3H), 3.73–3.66 (m, 4H), 3.24 (dd, $J = 13.0, 6.9$ Hz, 2H), 2.81–2.74 (m, 4H), 1.58–1.51 (m, 2H), 0.92 (t, $J = 7.4$ Hz, 3H); HRMS (ESI) m/z found 402.2203 ($M+H^+$), while $C_{22}H_{28}FN_3O_3$ ($M+H^+$) requires 402.2187.

4.2.14. 5-((3-fluoro-4-methoxybenzyl)amino)-N-isobutyl-2-morpholinobenzamide (C5)

Yellow solid; 1H NMR (600 MHz, DMSO- d_6) δ 10.12 (t, 1H), 7.19 (d, $J = 2.9$ Hz, 1H), 7.16–7.08 (m, 4H), 6.63 (dd, $J = 8.7, 2.9$ Hz, 1H), 6.34 (t, $J = 6.1$ Hz, 1H), 4.19 (d, $J = 6.1$ Hz, 2H), 3.79 (s, 3H), 3.74–3.66 (m, 4H), 3.13 (t, $J = 6.4$ Hz, 2H), 2.81–2.75 (m, 4H), 1.85–1.77 (m, 1H), 0.91 (d, $J = 6.7$ Hz, 6H); HRMS (ESI) m/z found 416.2353 ($M+H^+$), while $C_{23}H_{30}FN_3O_3$ ($M+H^+$) requires 416.2344.

4.2.15. N-(cyclopropylmethyl)-5-((3-fluoro-4-methoxybenzyl)amino)-2-morpholinobenzamide (C6)

White solid; 1H NMR (600 MHz, Chloroform- d) δ 9.88 (s, 1H), 7.78 (d, $J = 2.8$ Hz, 1H), 7.18 (d, $J = 8.8$ Hz, 1H), 7.10–7.01 (m, 2H), 6.91 (t, $J = 8.4$ Hz, 1H), 6.69 (dd, $J = 8.9, 2.6$ Hz, 1H), 4.31 (s, 2H), 4.02 (t, $J = 4.5$ Hz, 4H), 3.86 (s, 3H), 3.36–3.24 (m, 6H), 1.19–1.06 (m, 1H), 0.63–0.51 (m, 2H), 0.38–0.25 (m, 2H); ^{13}C NMR (150 MHz, Chloroform- d) δ 165.94, 152.45 (d, $J = 246.1$ Hz), 146.71 (d, $J = 10.7$ Hz), 145.44, 141.19, 132.21 (d, $J = 5.5$ Hz), 128.63, 123.01 (d, $J = 3.5$ Hz), 122.49, 115.69, 115.58, 115.14 (d, $J = 18.5$ Hz), 113.48 (d, $J = 2.2$ Hz), 67.52, 56.30, 53.77, 47.42, 44.61, 10.79, 3.92; HRMS (ESI) m/z found 414.2200 ($M+H^+$), while $C_{23}H_{28}FN_3O_3$ ($M+H^+$) requires 414.2187.

4.2.16. 5-((3-fluoro-4-methoxybenzyl)amino)-2-morpholino-N-neopentylbenzamide (C7)

Light yellow solid; 1H NMR (600 MHz, DMSO- d_6) δ 10.13 (t, $J = 6.4$ Hz, 1H), 7.22 (d, $J = 3.0$ Hz, 1H), 7.18–7.07 (m, 4H), 6.63 (dd, $J = 8.7, 3.0$ Hz, 1H), 6.35 (t, $J = 6.2$ Hz, 1H), 4.19 (d, $J = 6.1$ Hz, 2H), 3.79 (s, 3H), 3.72–3.67 (m, 4H), 3.15 (d, $J = 6.4$ Hz, 2H), 2.84–2.77 (m, 4H), 0.91 (s, 9H); HRMS (ESI) m/z found 430.2508 ($M+H^+$), while $C_{24}H_{32}FN_3O_3$ ($M+H^+$) requires 430.2500.

4.2.17. N-(cyclobutylmethyl)-5-((3-fluoro-4-methoxybenzyl)amino)-2-morpholinobenzamide (C8)

White solid; 1H NMR (300 MHz, Chloroform- d) δ 10.32 (s, 1H), 7.58 (d, $J = 2.9$ Hz, 1H), 7.11–7.01 (m, 3H), 6.88 (t, $J = 8.4$ Hz, 1H), 6.62 (dd, $J = 8.6, 2.9$ Hz, 1H), 4.26 (s, 2H), 3.91–3.74 (m, 7H), 3.47 (dd, $J = 7.4, 5.5$ Hz, 2H), 3.01–2.82 (m, 4H), 2.69–2.46 (m, 1H), 2.18–2.00 (m, 2H), 1.98–1.85 (m, $J = 7.3, 6.6$ Hz, 2H), 1.83–1.62 (m, 2H); HRMS (ESI) m/z found 428.2345 ($M+H^+$), while $C_{24}H_{30}FN_3O_3$ ($M+H^+$) requires 428.2344.

4.2.18. N-(cyclopentylmethyl)-5-((3-fluoro-4-methoxybenzyl)amino)-2-morpholinobenzamide (C9)

White solid; 1H NMR (600 MHz, DMSO- d_6) δ 10.10 (t, $J = 5.7$ Hz, 1H), 7.19 (d, $J = 2.9$ Hz, 1H), 7.17–7.13 (m, 1H), 7.12–7.07 (m, 3H), 6.63 (dd, $J = 8.7, 3.0$ Hz, 1H), 6.34 (t, $J = 6.2$ Hz, 1H), 4.19 (d, $J = 6.1$ Hz, 2H), 3.79 (s, 3H), 3.72–3.68 (m, 4H), 3.23 (dd, $J = 7.1, 5.9$ Hz, 2H), 2.80–2.76 (m, 4H), 2.14–2.08 (m, 1H), 1.75–1.67 (m, 2H), 1.62–1.55 (m, 2H), 1.54–1.47 (m, 2H), 1.24–1.21 (m, 2H); HRMS (ESI) m/z found 442.36 ($M+H^+$), while $C_{25}H_{32}FN_3O_3$ ($M+H^+$) requires 442.2500.

4.2.19. N-(cyclohexylmethyl)-5-((3-fluoro-4-methoxybenzyl)amino)-2-morpholinobenzamide (C10)

White solid; 1H NMR (600 MHz, DMSO- d_6) δ 10.16 (d, $J = 7.7$ Hz, 1H), 7.21 (d, $J = 3.0$ Hz, 1H), 7.17–7.08 (m, 4H), 6.62 (dd, $J = 8.7, 3.0$ Hz, 1H), 6.36 (t, $J = 6.2$ Hz, 1H), 4.19 (d, $J = 6.1$ Hz, 2H), 3.79 (s, 3H), 3.75–3.65 (m, 5H), 2.85–2.72 (m, 4H), 1.96–1.86 (m, 2H), 1.77–1.68 (m, 2H), 1.63–1.55 (m, 1H), 1.38–1.16 (m, 5H); HRMS (ESI) m/z found 442.2502 ($M+H^+$), while $C_{25}H_{32}FN_3O_3$ ($M+H^+$) requires 442.2500.

4.2.20. 5-((3-fluoro-4-methoxybenzyl)amino)-2-morpholinophenyl(piperidin-1-yl)methanone (C11)

Yellow solid; 1H NMR (600 MHz, DMSO) δ 7.15 (d, $J = 12.5$ Hz, 1H), 7.13–7.04 (m, 2H), 6.87 (d, $J = 8.7$ Hz, 1H), 6.53 (dd, $J = 8.7, 2.4$ Hz, 1H), 6.32 (s, 1H), 6.13 (s, 1H), 4.16 (s, 2H), 3.79 (s, 3H), 3.68–3.46 (m, 6H), 3.14–2.91 (m, 4H), 2.54 (s, 2H), 1.59–1.25 (m, 6H); HRMS (ESI) m/z found 428.2352 ($M+H^+$), while $C_{24}H_{30}FN_3O_3$ ($M+H^+$) requires 428.2344.

4.2.21. N-(2-fluoro-2-methylpropyl)-5-((3-fluoro-4-methoxybenzyl)amino)-2-morpholinobenzamide (C12)

Yellow solid; 1H NMR (300 MHz, Chloroform- d) δ 11.05 (s, 1H), 7.63 (d, $J = 3.0$ Hz, 1H), 7.19–7.02 (m, 3H), 6.90 (t, $J = 8.4$ Hz, 1H), 6.71 (dd, $J = 8.7, 2.8$ Hz, 1H), 4.29 (s, 2H), 3.89–3.81 (m, 7H), 3.64 (dd, $J = 22.9, 5.7$ Hz, 2H), 3.01–2.85 (m, 4H), 1.44 (s, 3H), 1.36 (s, 3H). HRMS (ESI) m/z found 434.2260 ($M+H^+$), While $C_{23}H_{29}F_2N_3O_3$ ($M+H^+$) requires 434.2177.

4.2.22. 5-((3-fluoro-4-methoxybenzyl)amino)-N-((1-hydroxycyclopropyl)methyl)-2-morpholinobenzamide (C13)

White solid; 1H NMR (600 MHz, Chloroform- d) δ 9.70 (s, 1H), 7.88 (s, 1H), 7.15 (d, $J = 8.4$ Hz, 1H), 6.97 (d, $J = 8.3$ Hz, 1H), 6.94–6.89 (m, 2H), 6.84 (t, $J = 8.4$ Hz, 1H), 4.79 (s, 2H), 3.94–3.90 (m, 4H), 3.85 (s, 3H), 3.82–3.80 (m, 2H), 3.06–3.00 (m, 4H), 2.02 (s, 3H), 1.02–0.97 (m, 2H), 0.97–0.91 (m, 2H). HRMS (ESI) m/z found 470.2081 ($M-H^-$), While $C_{25}H_{30}FN_3O_5$ ($M+H^+$) requires 472.2169.

4.2.23. 5-((3-fluoro-4-methoxybenzyl)amino)-N-(2-hydroxy-2-methylpropyl)-2-morpholinobenzamide (C14)

White solid; 1H NMR (600 MHz, Chloroform- d) δ 9.95 (s, 1H), 7.96 (s, 1H), 7.17 (d, $J = 8.4$ Hz, 1H), 6.97 (d, $J = 8.5$ Hz, 1H), 6.94–6.89 (m, 2H), 6.84 (t, $J = 8.5$ Hz, 1H), 4.80 (s, 2H), 3.93–3.89 (m, 4H), 3.85 (s, 3H), 3.49 (d, $J = 5.6$ Hz, 2H), 3.03–2.98 (m, 4H), 1.88 (s, 3H), 1.30 (s, 6H). HRMS (ESI) m/z found 474.2417 ($M+H^+$), While

$C_{25}H_{32}FN_3O_5(M+H)^+$ requires 474.2326.

4.2.24. *N*-(cyclopropylmethyl)-3-(3-fluoro-4-methoxybenzylamino)benzamide (**C15**)

White solid; 1H NMR (600 MHz, DMSO- d_6) δ 8.33 (t, $J = 5.8$ Hz, 1H), 7.20–7.16 (m, 1H), 7.15–7.06 (m, 4H), 7.01 (d, $J = 7.6$ Hz, 1H), 6.69 (dd, $J = 8.1, 2.4$ Hz, 1H), 6.39 (t, $J = 6.1$ Hz, 1H), 4.25 (d, $J = 6.0$ Hz, 2H), 3.80 (s, 3H), 3.11 (t, $J = 6.2$ Hz, 2H), 1.08–0.95 (m, 1H), 0.46–0.33 (m, 2H), 0.21 (t, $J = 4.7$ Hz, 2H); HRMS (ESI) m/z found 329.1673 ($M+H^+$), while $C_{19}H_{21}FN_2O_2$ ($M+H^+$) requires 329.1600.

4.2.25. *N*-(cyclopropylmethyl)-5-((3-fluoro-4-methoxybenzyl)amino)-2-(piperidin-1-yl)benzamide (**C16**)

White solid; 1H NMR (600 MHz, Chloroform- d) δ 11.04 (t, $J = 5.3$ Hz, 1H), 7.59 (d, $J = 3.0$ Hz, 1H), 7.15–7.03 (m, 3H), 6.92 (q, $J = 8.1$ Hz, 1H), 6.63 (dd, $J = 8.6, 3.0$ Hz, 1H), 4.27 (s, 2H), 3.88 (s, 3H), 3.31 (dd, $J = 7.2, 5.1$ Hz, 2H), 2.85 (t, $J = 5.2$ Hz, 4H), 1.81–1.70 (m, 4H), 1.14–1.00 (m, 1H), 0.64–0.50 (m, 2H), 0.35–0.24 (m, 2H). HRMS (ESI) m/z found 412.2396 ($M+H^+$); while $C_{24}H_{30}FN_3O_2$ ($M+H^+$) requires 412.2395.

4.2.26. *N*-(cyclopropylmethyl)-5-((3-fluoro-4-methoxybenzyl)amino)-2-(4-fluoropiperidin-1-yl)benzamide (**C17**)

White solid; 1H NMR (600 MHz, DMSO- d_6) δ 10.40 (t, $J = 5.3$ Hz, 1H), 7.23 (d, $J = 3.0$ Hz, 1H), 7.16–7.12 (m, 1H), 7.11–7.06 (m, 3H), 6.62 (dd, $J = 8.6, 3.0$ Hz, 1H), 6.33 (t, $J = 6.2$ Hz, 1H), 4.89–4.75 (m, 1H), 4.18 (d, $J = 6.1$ Hz, 2H), 3.78 (s, 3H), 3.14 (dd, $J = 7.1, 5.3$ Hz, 2H), 2.91 (t, $J = 9.6$ Hz, 2H), 2.76–2.71 (m, 2H), 2.06–1.96 (m, 2H), 1.94–1.84 (m, 2H), 1.04–0.97 (m, 1H), 0.49–0.45 (m, 2H), 0.24–0.21 (m, 2H). HRMS (ESI) m/z found 430.2306 ($M+H^+$); while $C_{24}H_{29}F_2N_3O_2$ ($M+H^+$) requires 430.2301.

4.2.27. *N*-(cyclopropylmethyl)-2-(4,4-difluoropiperidin-1-yl)-5-((3-fluoro-4-methoxybenzyl)amino)benzamide (**C18**)

White solid; 1H NMR (600 MHz, DMSO- d_6) δ 9.90 (t, $J = 5.4$ Hz, 1H), 7.18–7.12 (m, 2H), 7.12–7.05 (m, 3H), 6.61 (dd, $J = 8.7, 3.0$ Hz, 1H), 6.35 (t, $J = 6.2$ Hz, 1H), 4.19 (d, $J = 6.1$ Hz, 2H), 3.79 (s, 3H), 3.15 (dd, $J = 7.1, 5.4$ Hz, 2H), 2.91 (t, $J = 5.7$ Hz, 4H), 2.21–2.07 (m, 4H), 1.08–0.97 (m, 1H), 0.50–0.41 (m, 2H), 0.30–0.19 (m, 2H). HRMS (ESI) m/z found 448.2214 ($M+H^+$); while $C_{24}H_{28}F_3N_3O_2$ ($M+H^+$) requires 448.2206.

4.2.28. *N*-(cyclopropylmethyl)-5-((3-fluoro-4-methoxybenzyl)amino)-2-(4-hydroxypiperidin-1-yl)benzamide (**C19**)

White solid; 1H NMR (300 MHz, Chloroform- d) δ 10.48 (s, 1H), 7.58 (s, 1H), 7.14–6.99 (m, 3H), 6.91 (t, $J = 8.4$ Hz, 1H), 6.73–6.56 (m, 1H), 4.28 (s, 2H), 3.87 (s, 3H), 3.59 (s, 4H), 3.30 (dd, $J = 7.1, 4.7$ Hz, 2H), 2.88 (s, 4H), 1.49 (s, 9H), 1.17–0.94 (m, 1H), 0.64–0.45 (m, 2H), 0.37–0.16 (m, 2H). HRMS (ESI) m/z found 513.2877 ($M+H^+$); while $C_{28}H_{37}FN_4O_4$ ($M+H^+$) requires 513.2872.

4.2.29. *N*-(cyclopropylmethyl)-5-((3-fluoro-4-methoxybenzyl)amino)-2-(tetrahydro-2H-pyran-4-yl)benzamide (**C20**)

White solid; 1H NMR (600 MHz, Chloroform- d) δ 7.11 (d, $J = 8.5$ Hz, 1H), 7.08–7.01 (m, 2H), 6.90 (t, $J = 8.4$ Hz, 1H), 6.63 (dd, $J = 8.6, 2.3$ Hz, 1H), 6.56 (d, $J = 2.5$ Hz, 1H), 5.89 (t, $J = 5.5$ Hz, 1H), 4.22 (s, 2H), 3.99 (dd, $J = 11.3, 4.1$ Hz, 2H), 3.86 (s, 3H), 3.47 (td, $J = 11.7, 2.4$ Hz, 2H), 3.26 (dd, $J = 7.1, 5.5$ Hz, 2H), 3.11–3.00 (m, 1H), 1.79–1.65 (m, 4H), 1.06–0.94 (m, 1H), 0.57–0.49 (m, 2H), 0.28–0.20 (m, 2H). HRMS (ESI) m/z found 413.2230 ($M+H^+$); while

$C_{24}H_{29}FN_2O_3$ ($M+H^+$) requires 413.2235.

4.2.30. *N*-(cyclopropylmethyl)-4-((3-fluoro-4-methoxybenzyl)amino)-[1,1'-biphenyl]-2-carboxamide (**C21**)

White solid; 1H NMR (600 MHz, Chloroform- d) δ 7.36 (d, $J = 3.6$ Hz, 4H), 7.29 (m, 1H), 7.17 (d, $J = 8.3$ Hz, 1H), 7.13–7.06 (m, 2H), 7.03 (d, $J = 2.5$ Hz, 1H), 6.92 (t, $J = 8.4$ Hz, 1H), 6.72 (dd, $J = 8.4, 2.6$ Hz, 1H), 5.27 (t, $J = 5.5$ Hz, 1H), 4.32 (s, 2H), 3.87 (s, 2H), 2.95 (dd, $J = 7.3, 5.3$ Hz, 2H), 0.62–0.48 (m, 1H), 0.30–0.20 (m, 2H), –0.07 to –0.18 (m, 2H). HRMS (ESI) m/z found 405.1975 ($M+H^+$); while $C_{25}H_{25}FN_2O_2$ ($M+H^+$) requires 405.1973.

4.2.31. *N*-(cyclopropylmethyl)-5-((3-fluoro-4-methoxybenzyl)(methyl)amino)-2-morpholinobenzamide (**C22**)

Light yellow solid; 1H NMR (600 MHz, DMSO- d_6) δ 10.24 (t, $J = 5.3$ Hz, 1H), 7.31 (d, $J = 3.2$ Hz, 1H), 7.21 (d, $J = 8.9$ Hz, 1H), 7.09 (t, $J = 8.7$ Hz, 1H), 7.02 (dd, $J = 12.4, 2.1$ Hz, 1H), 6.98–6.93 (m, 1H), 6.83 (dd, $J = 8.9, 3.3$ Hz, 1H), 4.48 (s, 2H), 3.79 (s, 3H), 3.77–3.74 (m, 4H), 3.16 (dd, $J = 7.1, 5.3$ Hz, 2H), 2.85–2.80 (m, 4H), 1.09–0.97 (m, 1H), 0.53–0.46 (m, 2H), 0.29–0.20 (m, 2H); HRMS (ESI) m/z found 428.2355 ($M+H^+$); while $C_{24}H_{30}FN_3O_3$ ($M+H^+$) requires 428.2344.

4.2.32. 5-(*N*-(3-fluoro-4-methoxybenzyl)acetamido)-*N*-((1-hydroxycyclopropyl)methyl)-2-morpholinobenzamide (**C23**)

White solid; 1H NMR (600 MHz, Chloroform- d) δ 9.70 (s, 1H), 7.88 (s, 1H), 7.15 (d, $J = 8.4$ Hz, 1H), 6.97 (d, $J = 8.3$ Hz, 1H), 6.94–6.89 (m, 2H), 6.84 (t, $J = 8.4$ Hz, 1H), 4.79 (s, 2H), 3.94–3.90 (m, 4H), 3.85 (s, 3H), 3.82–3.80 (m, 2H), 3.06–3.00 (m, 4H), 2.02 (s, 3H), 1.02–0.97 (m, 2H), 0.97–0.91 (m, 2H). HRMS (ESI) m/z found 470.2081 ($M-H^-$), While $C_{25}H_{30}FN_3O_5$ ($M+H^+$) requires 472.2169.

4.2.33. 5-(*N*-(3-fluoro-4-methoxybenzyl)acetamido)-*N*-(2-hydroxy-2-methylpropyl)-2-morpholinobenzamide (**C24**)

White solid; 1H NMR (600 MHz, Chloroform- d) δ 9.95 (s, 1H), 7.96 (s, 1H), 7.17 (d, $J = 8.4$ Hz, 1H), 6.97 (d, $J = 8.5$ Hz, 1H), 6.94–6.89 (m, 2H), 6.84 (t, $J = 8.5$ Hz, 1H), 4.80 (s, 2H), 3.93–3.89 (m, 4H), 3.85 (s, 3H), 3.49 (d, $J = 5.6$ Hz, 2H), 3.03–2.98 (m, 4H), 1.88 (s, 3H), 1.30 (s, 6H). HRMS (ESI) m/z found 474.2417 ($M+H^+$), While $C_{25}H_{32}FN_3O_5$ ($M+H^+$) requires 474.2326.

4.2.34. *N*-(cyclopropylmethyl)-5-(3-fluoro-4-methoxybenzylamido)-2-morpholinobenzamide (**C25**)

White solid; 1H NMR (600 MHz, DMSO- d_6) δ 10.34 (t, $J = 5.3$ Hz, 1H), 7.22 (d, $J = 2.9$ Hz, 1H), 7.08 (d, $J = 8.5$ Hz, 1H), 6.65 (dd, $J = 8.5, 2.9$ Hz, 1H), 5.12 (s, 2H), 3.75 (dd, $J = 5.6, 3.6$ Hz, 4H), 3.16 (dd, $J = 7.0, 5.3$ Hz, 2H), 2.80 (t, $J = 4.6$ Hz, 4H), 1.07–1.00 (m, 1H), 0.52–0.47 (m, 2H), 0.28–0.23 (m, 2H). HRMS (ESI) m/z found 428.1986 ($M+H^+$), While $C_{23}H_{26}FN_3O_4$ ($M+H^+$) requires 428.1907.

4.3. Biology

4.3.1. Assay for ALDH2 activation

The ALDH2 test kit (ab115348) and active human ALDH2 full length protein (ab87415) were purchased from Abcam (Cambridge, United Kingdom), and the assay was performed by BioTek Synergy 2 Multi-Mode Microplate Readers (Winooski, Vermont, United States of America). ALDH2 can oxidize acetaldehydes to acetic acids with the aid of NAD^+ while reducing NAD^+ to NADH. Then, NADH production was coupled with the dye to form a yellow product and the absorbance value of the yellow product could be monitored at an excitation wavelength of 370 nm and an emission wavelength of

450 nm. The assay is conducted using 5 µg/ml, 25 mM and 1 mM as the starting concentration of ALDH2, acetaldehyde and NAD⁺ under room temperature. The activity of the ALDH2 can be expressed as the change in absorbance per minute (SlopeOD450). The maximum fold activation was calculated as the activity of the ALDH2 with target compound at 100 µM treatment divided by that without treatment ($\text{SlopeOD450}_{\text{compound}}/\text{SlopeOD450}_{\text{min}}$). After corrected by the activities of Alda-1 ($\text{SlopeOD450}_{\text{max}}$) as 100% [$(\text{SlopeOD450}_{\text{compound}} - \text{SlopeOD450}_{\text{min}})/(\text{SlopeOD450}_{\text{max}} - \text{SlopeOD450}_{\text{min}}) * 100$], it was taken as the primary indicator. The EC₅₀ and standard error of selected compounds were determined from the activity of the ALDH2-dose curves by Graphpad Prism (GraphPad Software, Inc., San Diego, California, United States of America), which was taken as the second indicator.

4.3.2. Assay for protective effect against ischemic stroke in MCAO rats

Animals were maintained in accordance with institutional guidelines as defined by Institutional Animal Care and Use Committee for U.S. institutions. Male SD rats were provided by Shanghai Experimental Animal Center of Chinese Academy of Sciences (Shanghai, China) were housed in controlled conditions and received a standard rat chow and tap water ad libitum. Male SD rats (200–250 g) were randomized into the 3 groups: dimethyl sulfoxide (DMSO) (n = 6); **C6** (n = 6, 64 µg in 5 µl). These rats were injected (i.c.v.) with **C6** or saline 5 min before transient middle cerebral artery occlusion (MCAO) operation.

Anesthesia was induced by 1% isoflurane in 30% oxygen and 70% nitrous oxide with a face mask. Ischemia was induced by placing an 11-mm silicone-coated 8–0 filament from the left common carotid artery into the internal carotid artery for 2 h. Rectal temperature was maintained at 37 °C ± 0.5 °C with a control unit (Frederick Haer and Co) during anesthesia, thereafter animals were maintained in an incubator at 31 °C. Following reperfusion 24 h, rats were re-anesthetized and brains were immediately removed. Infarct size was measured by the TTC staining assay. Briefly, the brains were cut into 1–1.5 mm thick slices perpendicular to the brain axis. The slices were put in 0.8% triphenyltetrazolium chloride (TTC) for 15 min at 37 °C to distinguish the viable brain from the necrotic. After 1 h of incubation in 4% formaldehyde, slices were scanned and the areas were determined by planimetry of computer images (Image-Pro Plus 6.0; Bethesda, MD, USA).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Shanghai Pujiang Program (2019PJJD058) and Shanghai Biological and Pharmaceutical Science and Technology Support Program (20S11903000).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ejmech.2020.113119>.

References

- [1] C.H. Chen, J.C. Ferreira, E.R. Gross, D. Mochly-Rosen, Targeting aldehyde dehydrogenase 2: new therapeutic opportunities, *Physiol. Rev.* 94 (2014) 1–34.
- [2] V. Koppaka, D.C. Thompson, Y. Chen, M. Ellermann, K.C. Nicolaou, R.O. Juvonen, D. Petersen, R.A. Deitrich, T.D. Hurley, V. Vasilou, Aldehyde dehydrogenase inhibitors: a comprehensive review of the pharmacology, mechanism of action, substrate specificity, and clinical application, *Pharm. Rev.* 64 (2012) 520–539.
- [3] E.R. Gross, V.O. Zambelli, B.A. Small, J.C. Ferreira, C.H. Chen, D. Mochly-Rosen, A personalized medicine approach for Asian Americans with the aldehyde dehydrogenase 2*2 variant, *Annu. Rev. Pharmacol. Toxicol.* 55 (2015) 107–127.
- [4] H.R. Luo, G.S. Wu, A.J. Pakstis, L. Tong, H. Oota, K.K. Kidd, Y.P. Zhang, Origin and dispersal of atypical aldehyde dehydrogenase ALDH2487Lys, *Gene* 435 (2009) 96–103.
- [5] H. Li, S. Borinskaya, K. Yoshimura, N. Kal'ina, A. Marusin, V.A. Stepanov, Z. Qin, S. Khaliq, M.Y. Lee, Y. Yang, A. Mohyuddin, D. Gurwitz, S.Q. Mehdi, E. Rogava, L. Jin, N.K. Yankovsky, J.R. Kidd, K.K. Kidd, Refined geographic distribution of the oriental ALDH2*504Lys (nee 487Lys) variant, *Ann. Hum. Genet.* 73 (2009) 335–345.
- [6] J.I. Garaycochea, G.P. Crossan, F. Langevin, L. Mulderrig, S. Louzada, F. Yang, G. Guilbaud, N. Park, S. Roerink, S. Nik-Zainal, M.R. Stratton, K.J. Patel, Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells, *Nature* 553 (2018) 171–177.
- [7] M. Matejic, M.J. Gunter, P. Ferrari, Alcohol metabolism and oesophageal cancer: a systematic review of the evidence, *Carcinogenesis* 38 (2017) 859–872.
- [8] C.H. Chen, L. Sun, D. Mochly-Rosen, Mitochondrial aldehyde dehydrogenase and cardiac diseases, *Cardiovasc. Res.* 88 (2010) 51–57.
- [9] Y. Mizuno, S. Hokimoto, E. Harada, K. Kinoshita, K. Nakagawa, M. Yoshimura, H. Ogawa, H. Yasue, Variant aldehyde dehydrogenase 2 (ALDH2*2) is a risk factor for coronary spasm and ST-segment elevation myocardial infarction, *J. Am. Heart Assoc.* 5 (2016), e003247.
- [10] X.J. Luo, B. Liu, Q.L. Ma, J. Peng, Mitochondrial aldehyde dehydrogenase, a potential drug target for protection of heart and brain from ischemia/reperfusion injury, *Curr. Drug Targets* 15 (2014) 948–955.
- [11] C.H. Chen, G.R. Budas, E.N. Churchill, M.H. Disatnik, T.D. Hurley, D. Mochly-Rosen, Activation of aldehyde dehydrogenase-2 reduces ischemic damage to the heart, *Science* 321 (2008) 1493–1495.
- [12] Y.F. Sung, C.C. Lu, J.T. Lee, Y.J. Hung, C.J. Hu, J.S. Jeng, H.Y. Chiou, G.S. Peng, Homozygous ALDH2*2 is an independent risk factor for ischemic stroke in Taiwanese men, *Stroke* 47 (2016) 2174–2179.
- [13] Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study, *Lancet Neurol.* 15 (2016) 174–184.
- [14] W. Wang, L.L. Lin, J.M. Guo, Y.Q. Cheng, J. Qian, J.L. Mehta, D.F. Su, P. Luan, A.J. Liu, Heavy ethanol consumption aggravates the ischemic cerebral injury by inhibiting ALDH2, *Int. J. Stroke* 10 (2015) 1261–1269.
- [15] J.M. Guo, A.J. Liu, P. Zang, W.Z. Dong, L. Ying, W. Wang, P. Xu, X.R. Song, J. Cai, S.Q. Zhang, J.L. Duan, J.L. Mehta, D.F. Su, ALDH2 protects against stroke by clearing 4-HNE, *Cell Res.* 23 (2013) 915–930.
- [16] S. Perez-Miller, H. Younus, R. Vanam, C.H. Chen, D. Mochly-Rosen, T.D. Hurley, Alda-1 is an agonist and chemical chaperone for the common human aldehyde dehydrogenase 2 variant, *Nat. Struct. Mol. Biol.* 17 (2010) 159–164.
- [17] H.N. Larson, J. Zhou, Z. Chen, J.S. Stampler, H. Weiner, T.D. Hurley, Structural and functional consequences of coenzyme binding to the inactive asian variant of mitochondrial aldehyde dehydrogenase: roles of residues 475 and 487, *J. Biol. Chem.* 282 (2007) 12940–12950.
- [18] H.N. Larson, H. Weiner, T.D. Hurley, Disruption of the coenzyme binding site and dimer interface revealed in the crystal structure of mitochondrial aldehyde dehydrogenase "Asian" variant, *J. Biol. Chem.* 280 (2005) 30550–30556.
- [19] J. Hu, W. Tian, R. Zhou, Y. Zhang, J. Lv, J. Zhu, X. Chen, X. Pan, C. Zheng, Design, synthesis, and biological evaluation of new ALDH2 activators, *J. Saudi Chem. Soc.* 23 (2019) 255–262.
- [20] C. Zhuang, S. Narayanapillai, W. Zhang, Y.Y. Sham, C. Xing, Rapid identification of Keap1-Nrf2 small-molecule inhibitors through structure-based virtual screening and hit-based substructure search, *J. Med. Chem.* 57 (2014) 1121–1126.
- [21] T. Zhu, S. Cao, P.C. Su, R. Patel, D. Shah, H.B. Chokshi, R. Szukala, M.E. Johnson, K.E. Hevener, Hit identification and optimization in virtual screening: practical recommendations based on a critical literature analysis, *J. Med. Chem.* 56 (2013) 6560–6572.
- [22] X. Liu, G. Dong, J. Zhang, J. Qi, C. Zheng, Y. Zhou, J. Zhu, C. Sheng, J. Lu, Discovery of novel human acrosin inhibitors by virtual screening, *J. Comput. Aided Mol. Des.* 25 (2011) 977–985.
- [23] G. Dong, C. Sheng, S. Wang, Z. Miao, J. Yao, W. Zhang, Selection of evodiamine as a novel topoisomerase I inhibitor by structure-based virtual screening and hit optimization of evodiamine derivatives as antitumor agents, *J. Med. Chem.* 53 (2010) 7521–7531.
- [24] A.R. Leach, V.J. Gillet, R.A. Lewis, R. Taylor, Three-dimensional pharmacophore methods in drug discovery, *J. Med. Chem.* 53 (2010) 539–558.
- [25] D. Pathak, N. Chadha, O. Silakari, Identification of non-resistant ROS-1 inhibitors using structure based pharmacophore analysis, *J. Mol. Graph. Model.* 70 (2016) 85–93.
- [26] J. Meslamani, J. Li, J. Sutter, A. Stevens, H.O. Bertrand, D. Rognan, Protein-ligand-based pharmacophores: generation and utility assessment in computational ligand profiling, *J. Chem. Inf. Model.* 52 (2012) 943–955.
- [27] Glide, Schrodinger, LLC: New York, NY, United States of America.
- [28] W. Yang, C.H. Chen, D. Mochly-Rosen, Preparation of benzamides and pyridinecarboxamides as mitochondrial aldehyde dehydrogenase-2 modulators,

- PCT Int. Appl. (2014), WO 2014160185 A2.
- [29] P.S. Campbell, C. Jamieson, I. Simpsonb, A.J.B. Watson, Practical synthesis of pharmaceutically relevant molecules enriched in sp^3 character, *Chem. Commun.* 54 (2018) 46–49.
- [30] Discovery Studio. Accelrys, Inc.: San Diego, California, United States of America.
- [31] GOLD. The Cambridge Crystallographic Data Centre: Cambridge, United Kingdom.
- [32] R. Thomsen, M.H. Christensen, MolDock: a new technique for high-accuracy molecular docking, *J. Med. Chem.* 49 (2006) 3315–3321.