

Discovery of a Natural-Product-Derived Preclinical Candidate for Once-Weekly Treatment of Type 2 Diabetes

Shiliang Li, Chun Qin, Shichao Cui, Hongling Xu, Fangshu Wu, Jiawei Wang, Mingbo Su, Xiaoyu Fang, Dan Li, Qian Jiao, Ming Zhang, Chunmei Xia, Lili Zhu, Rui Wang, Jia Li, Hualiang Jiang, Zhenjiang Zhao, Jing-Ya Li, and Honglin Li

J. Med. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jmedchem.8b01491 • Publication Date (Web): 29 Jan 2019

Downloaded from <http://pubs.acs.org> on January 30, 2019

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Discovery of a Natural-Product-Derived Preclinical Candidate for Once-Weekly Treatment of Type 2 Diabetes

Shiliang Li^{1,†}, Chun Qin^{1,†}, Shichao Cui^{2,3,†}, Hongling Xu¹, Fangshu Wu¹, Jiawei Wang¹, Mingbo Su², Xiaoyu Fang¹, Dan Li², Qian Jiao¹, Ming Zhang¹, Chunmei Xia², Lili Zhu¹, Rui Wang¹, Jia Li², Hualiang Jiang², Zhenjiang Zhao^{1,*}, Jingya Li^{2,3,*}, Honglin Li^{1,*}

¹ *Shanghai Key Laboratory of New Drug Design, State Key Laboratory of Bioreactor Engineering, School of Pharmacy, East China University of Science and Technology, Shanghai 200237, China.*

² *State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences. Shanghai 201203, China.*

³ *University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing, 100049, P. R. China*

† Authors contributed equally to this work

*To whom correspondence should be addressed. Email: zhjzhao@ecust.edu.cn, jyli@simm.ac.cn, hlli@ecust.edu.cn

Abstract

Poor medication adherence is one of the leading causes of suboptimal glycaemic control in approximately half of the patients with type 2 diabetes mellitus (T2DM). Long-acting antidiabetic drugs are clinically needed for improving patients' compliance. DPP-4 inhibitors play an increasingly important role in the treatment of T2DM for their favorable properties of weight neutrality and hypoglycemia avoidance. Herein, we report the successful discovery and scale-up synthesis of compound **5**, a structurally novel, potent and long-acting DPP-4 inhibitor for the once-weekly treatment of T2DM. Inhibitor **5** has fast-associating and slow-dissociating binding kinetics profiles as well as slow clearance rate and long terminal half-life pharmacokinetic properties. A single dose oral administration of **5** (3 mg/kg) inhibited >80% DPP-4 activity for more than 7 days in diabetic mice. The long-term antidiabetic efficacies of **5** (10 mg/kg, q.w) were better than the once-weekly trelagliptin and omarigliptin, especially in decreasing the hemoglobin A1c (HbA1c) level.

Keywords: Drug discovery, DPP-4 inhibitor, Natural product, Diabetes, Once-weekly treatment

Introduction

Diabetes mellitus is one of the largest health emergencies globally. The global healthcare costs to treat diabetes and related complications have reached 727 billion dollars in 2017.¹ The worldwide prevalence of diabetes mellitus for adults is growing at a continuous rate and has reached a total of 425 million in 2017, and the number will rise to 629 million by 2045. Among the diagnosed patients, approximately 90% of cases are type 2 diabetes mellitus (T2DM). Diabetes causes significant morbidity and mortality. Prevention of diabetes-related micro- and macrovascular complications such as nephropathy, retinopathy, neuropathy, and stroke in diabetic patients requires tight glycemic control. An increasing number of antihyperglycemic agents have been approved for the treatment of T2DM. Compared with the traditional antidiabetic drugs (mainly biguanides, sulfonylureas, thiazolidinediones, meglitinides, α -glucosidase inhibitors and insulin) that usually bring undesirable side effects including weight gain, hypoglycemia and gastrointestinal distress, newer therapies such as glucagon-like peptide-1 (GLP-1) receptor agonists, DPP-4 inhibitors and sodium-glucose co-transporter 2 (SGLT2) inhibitors have more advantages because of their efficacy and safety.² Nevertheless, in spite of their potency in lowering HbA1c level without causing hypoglycemia, a particular inconvenience with GLP-1 receptor agonists therapy is its administration by injection, while SGLT2 inhibitors may potentially lower blood pressure and increase genital infection. Thus, DPP-4 inhibitors tend to draw more attention.^{3,4}

The DPP-4 enzyme is a widely expressed serine protease that belongs to the α/β -hydrolases (family S9B). It can selectively inactivate peptides with an alanine, proline, or

1
2
3
4 serine residue at the penultimate position from the N-terminus.⁵ Inhibition of the catalytic
5
6 activity of DPP-4 helps to elevate GLP-1 level, which is an important incretin secreted by the
7
8 small intestine in response to meal ingestion, and amplifies the insulin secretory response to
9
10 nutrients through their cognate G protein-coupled receptors in the pancreatic β cells.⁶ GLP-1
11
12 regulates glucose lowering activities in a glucose-dependent manner, however, it has short
13
14 half-lives *in vivo* (< 2 min), and is rapidly degraded to their inactive form by DPP-4.⁶
15
16 Therefore, by inhibiting the catalytic activity of DPP-4, DPP-4 inhibitor helps to maintain the
17
18 endogenous GLP-1 concentration, and is a clinically proven, therapeutically novel approach
19
20 to treat T2DM.⁷
21
22
23
24
25

26
27 Nowadays, eleven DPP-4 inhibitors have been approved worldwide (Fig. S1). Classical
28
29 sitagliptin (Fig. 1a), saxagliptin, alogliptin, and linagliptin et al. are all DPP-4 inhibitors given
30
31 once per day. Since T2DM is a chronic, lifelong condition that requires patients to take
32
33 anti-diabetic drugs regularly to control glucose levels, long-acting DPP-4 inhibitors have
34
35 been clinically needed for improving patients' compliance.⁸ Trelagliptin and omarigliptin are
36
37 the only two gliptins that have recently received marketing authorization in Japan in 2015 for
38
39 once-weekly dosing regimen.^{9, 10} Patient satisfaction and compliance improved after
40
41 switching from taking once-daily DPP-4 inhibitor to once-weekly DPP-4 inhibitor in Japan.^{11,}
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
12 However, manufacturers of both drugs have announced that they discontinued their plans
for seeking regulatory approval in countries other than Japan. Omarigliptin was reported to
have a higher rate of serious adverse events than sitagliptin, showing elevated rates of
cardiovascular issues, hepatobiliary events and prostate cancer.¹³⁻¹⁶ Although currently
available data of trelagliptin suggest that trelagliptin is a safe and efficacious medication with

1
2
3
4 a similar safety profile to the once-daily alogliptin,¹⁷ further studies are needed to investigate
5
6 the effect of trelagliptin on cardiovascular events,¹⁸ since its close analog alogliptin had been
7
8 warned by the U.S. FDA for an increased risk of heart failure.¹⁹ It is still unclear whether the
9
10 excessive heart failure effect is target related or compound related, because sitagliptin did not
11
12 increase cardiovascular risk in T2DM patients.^{20, 21} Structurally distinct DPP-4 inhibitors may
13
14 serve as new tools to address this question. Therefore, it is still a necessity and challenge to
15
16 find novel and safe DPP-4 inhibitors with long-acting antidiabetic efficacy, especially in
17
18 China where more than 110 million adults have T2DM, accounting for about 25% of all
19
20 T2DM patients worldwide.
21
22
23
24
25

26
27 Previously, based on a natural product lead *iso*-daphnetin (Fig. 1a), we have discovered
28
29 a novel series of 3-phenyl-2,3-dihydro-1*H*-benzo[*f*]chromen-2-amine derivatives as potent
30
31 DPP-4 inhibitors by scaffold hopping.²² Among those inhibitors, compound **1** (Fig. 1a) was
32
33 the most promising one, which could inhibit > 80% DPP-4 activity for more than 24 h with a
34
35 single oral dose of 3 mg/kg in diabetic mice. However, the long-acting antidiabetic efficacy
36
37 of **1** still cannot afford a once-weekly dosing regimen. Both the ligand-protein binding
38
39 kinetics and pharmacokinetic profiles of **1** should be optimized because those two profiles are
40
41 reported to directly influence the duration of a drug's *in vivo* efficacy.²³⁻²⁶
42
43
44
45
46
47

48 Herein, based on the lead compound **1**, we report the successful discovery, scale-up
49
50 synthesis and biochemical, structural, pharmacokinetic as well as pharmacological evaluation
51
52 of a natural-product-derived DPP-4 inhibitor **5** for once-weekly treatment of T2DM with
53
54 better and longer antidiabetic efficacy than trelagliptin and omarigliptin.
55
56
57
58
59
60

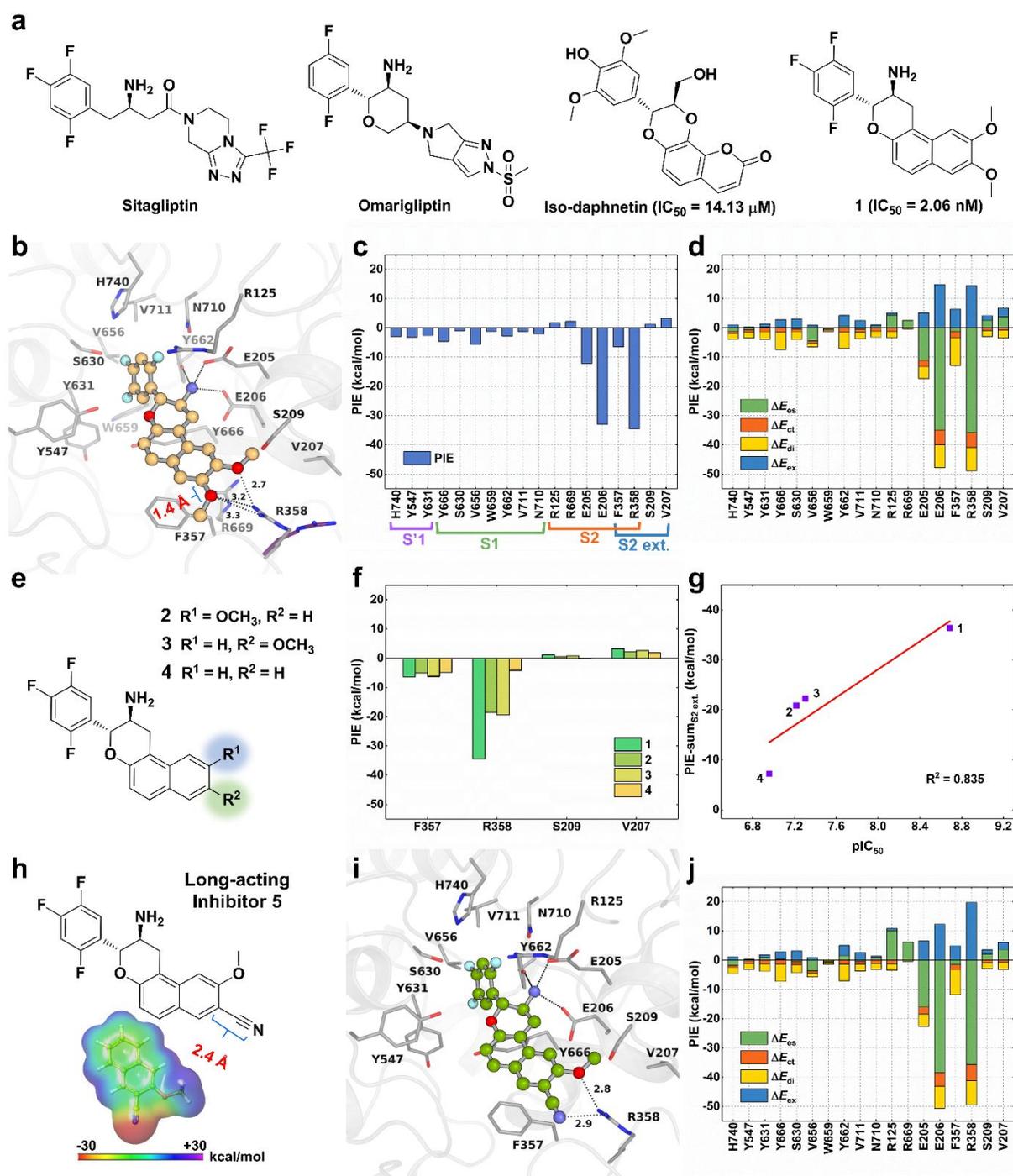


Fig. 1 Discovery and design of long-acting DPP-4 inhibitor **5**. **(a)** Chemical structures of representative DPP-4 inhibitors. **(b)** Binding mode of DPP-4 inhibitor **1** (PDB code: 5J3J), key residues are shown in grey sticks, while Arg358 from 1X70 binding sitagliptin as a ligand is shown in purple for comparison. Hydrogen bonds are displayed as black dashes. **(c)** Pair interaction energy (PIE) between the residues at the active site of DPP-4 and inhibitor **1**

1
2
3
4 and (d) Pair interaction energy decomposition analysis (PIEDA) for these residues. The
5
6 electrostatics, charge-transfer, dispersion, and exchange-repulsion PIE terms are represented
7
8 in green, orange, yellow, and blue, respectively. (e) Structures of three close analogs of **1**:
9
10 compounds **2-4**. (f) PIEs between residues at the S2 extended subsite of DPP-4 and **1-4**. (g)
11
12 Correlations between experimentally measured affinities (DPP-4, pIC₅₀) and sums of PIEs
13
14 between residues at the S2 extended subsite (PIE-sum_{S2 ext.}) of DPP-4 and inhibitors **1-4**. (h)
15
16 Structure of long-acting candidate **5** and the electrostatic surface potential (ESP) of its
17
18 fragment of 8-cyano-9-methoxy-naphthalene. (i) Simulated binding mode of **5** with Arg358
19
20 in a closed-up conformation for FMO calculation. (j) PIEDA for the residues at the active site
21
22 of DPP-4 and inhibitor **5**.
23
24
25
26
27
28
29
30
31
32

33 Results

34 Design of the potent and long-acting DPP-4 inhibitor

35
36
37
38
39 In our previous study, we found that both shape and electrostatic complementarities at
40
41 the S2 extended subsite of the DPP-4 catalytic site influence the bioactivity of the inhibitors
42
43 in our series.²² The flexible Arg358 can be attracted by the negative electrostatic potential
44
45 (ESP_{min} = -36.41 kcal/mol) produced by the two methoxyl oxygen atoms in **1**, and rotated its
46
47 guanidine group into a closed-up conformation to generate hydrogen bonds with the
48
49 methoxyl oxygen atoms (Fig. 1b). Therefore, compound **1** has a nanomolar binding affinity
50
51 ($K_D = 2.38 \times 10^{-9}$ M) against DPP-4 and a slower dissociation rate ($k_{\text{off}} = 3.43 \times 10^{-3}$ s⁻¹) than
52
53 sitagliptin ($k_{\text{off}} = 1.49 \times 10^{-2}$ s⁻¹) whose Arg358 flips away in an open conformation.
54
55
56
57
58
59
60 Conformational changes were reported to be determinative features of many protein-ligand

1
2
3
4 systems with slow dissociation kinetics.²⁷ Thus, we speculated that interacting with Arg358
5
6 may be a key factor influencing the dissociation rate of our compounds in this series since
7
8
9 Arg358 is flexible like a door which can hinder the ligand from leaving.

10
11
12 In order to investigate how Arg358 contributes to the ligand-protein binding, the
13
14 fragment molecular orbital (FMO) quantum-mechanical (QM) method²⁸⁻³¹ was applied to
15
16 quantify the interactions formed between **1** and the pocket-lining residues of DPP-4. After
17
18 FMO calculation, we found that dispersion interactions are prominent at the S1 and S'1
19
20 subsites (Fig. 1c and 1d), indicating that hydrophobic interactions are major inter-molecular
21
22 interactions at those regions including residues His740, Tyr547, Try631, Tyr666, Val656,
23
24 Tyr662, Val711, Asn710. Strong electrostatic interactions were presented at the S2 subsite,
25
26 especially between **1** and residues Glu205 and Glu206, which are the two most important
27
28 residues in the binding pocket that most of the reported DPP-4 inhibitors are interacting with.
29
30 At the S2 extended subsite including residues Val207, Ser209, Phe357 and Arg358, we were
31
32 surprised to observe that apart from the strong π - π stacking interactions between **1** and
33
34 Phe357 informed by the pair interaction energy decomposition analysis (PIEDA), pair
35
36 interaction energy (PIE) between **1** and Arg358 is comparable to that of Glu206 (-34.49 vs
37
38 -32.94 kcal/mol), indicating that Arg358 contributes significantly to the binding stability of
39
40 the complex. To explore whether such a great contribution made by Arg358 to the binding
41
42 affinity is unique to the scaffold of **1**, we subsequently performed FMO calculation on the
43
44 X-ray complex structure of sitagliptin and DPP-4 (PDB: 1X70) in which Arg358 adopts an
45
46 open conformation to accommodate the -CF₃ group of sitagliptin. Results showed that PIE
47
48 between sitagliptin and Arg358 is -9.59 kcal/mol (Fig. S2), only a third of that of the
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 **1**-Arg358 pair. The above results suggested that the ligand-induced conformation change of
5
6 Arg358 plays an important role in the binding of **1** to DPP-4, and contributes significantly to
7
8 the binding affinity.
9

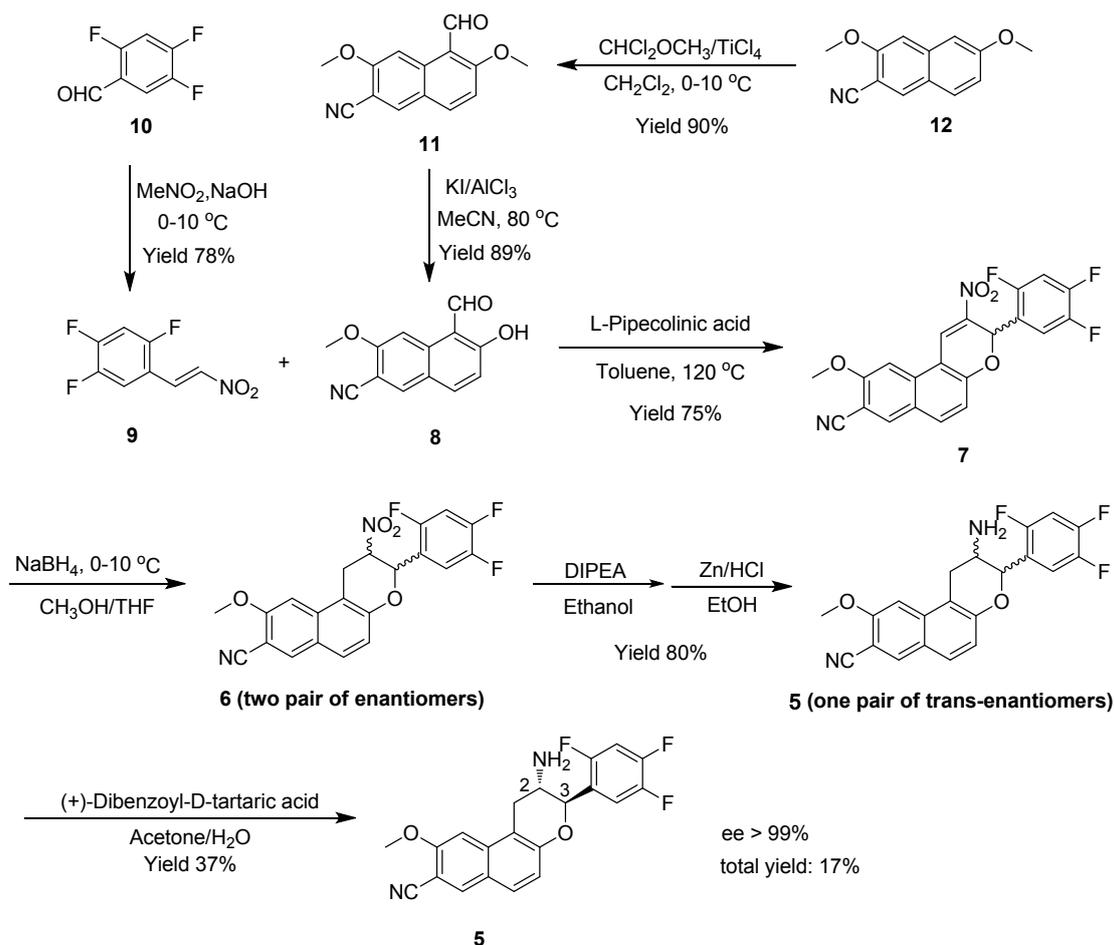
10
11
12 Expecting to keep Arg358 in a closed-up conformation in the subsequent structural
13
14 optimization, the specific contribution made by the two methoxyl groups should be studied to
15
16 illuminate us which methoxyl group is appropriate for substituting at the S2 extended subsite.
17
18 Then, FMO calculations were performed on three close analogs of **1** (Fig. 1e): compound **2**
19
20 with only one methoxyl group at R¹, compound **3** with only one methoxyl group at R², and
21
22 compound **4** with neither of the methoxyl groups. Preparations of the complex structures of
23
24 **2-4** with DPP-4 were described in the experimental section. For the four residues at the S2
25
26 extended subsite, major PIE differences were observed against Arg358 (Fig. 1f and Fig. S2).
27
28 We found that methoxyl group at R¹ and methoxyl group at R² play an almost equally
29
30 important role in interacting with Arg358, because PIEs of the **2**-Arg358 and **3**-Arg358 pairs
31
32 were close to each other (-18.52 vs -19.40 kcal/mol). Without two methoxyl groups, **4**
33
34 interacted Arg358 with PIE value of -4.16 kcal/mol, a more than 4 fold decrease than that of
35
36 **2** or **3**. Whereas, the occurrence of the two methoxyl group brought **1** the most strong
37
38 interaction with Arg358 (PIE = -34.49 kcal/mol). The sums of the PIEs (PIE-sum_{S2 ext.})
39
40 between each analog and residues at the S2 extended subsite correlated well with the four
41
42 inhibitors' biological potencies ($R^2 = 0.84$, Fig. 1g). Those results suggested that functional
43
44 groups at R¹ and R² should cooperate with each other to attract Arg358 to form a closed-up
45
46 conformation, and two methoxyl groups in **1** made nearly equal contributions. Compared
47
48 with R¹ position that has limited space, R² position is open wide enough, thus we decided to
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 decorate at R² with the attempt to keep Arg358's closed-up conformation.
5

6 The methoxyl oxygen atom of R² in **1** that makes H-bond with Arg358 is 1.4 Å away
7
8 from the attached aromatic carbon atom. In order to mimic the role of that oxygen atom, an
9
10 H-bond acceptor that attached close to the aromatic ring but with greatly increased
11
12 electro-withdrawing profile is preferred. A linear and rigid cyano group, which is 2.4 Å long
13
14 with an electronegative nitrogen atom that has a set of lone pair electrons, emerged in our
15
16 mind immediately. The ESP_{min} of the fragment containing R¹ of -OCH₃ and R² of -CN was
17
18 more negative than that of the fragment in **1** (-44.17 vs -36.41 kcal/mol) (Fig. 1h), indicating
19
20 that **5** may be capable to fix Arg358 in a closed-up conformation by strong dipole-dipole
21
22 interaction. We then constructed the complex structure of **5** and DPP-4 (Fig. 1i) and
23
24 performed the FMO calculation. PIEDA results informed us that the overall energy of the
25
26 electronic, charge transfer and dispersion interactions of the **5**-Arg358 pair (-49.45 kcal/mol)
27
28 is comparable to that of the **1**-Arg358 pair (-48.90 kcal/mol). Meanwhile, the introduction of
29
30 the cyano group made the molecule more polar in nature, leading to strengthened electrostatic
31
32 interactions between the amino group and residues Glu205 and Glu206 at the S2 subsite (Fig.
33
34 1j). Those calculating data suggested that **5** is of great possibility to lock Arg358 in a
35
36 closed-up conformation. Additionally, substituting -OCH₃ with -CN may also improve the
37
38 pharmacokinetic features of the compound. Subsequently, compound **5** was synthesized for
39
40 further experimental verification.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55

56 **Scale-up synthesis of compound 5**

57
58
59
60



Scheme 1. Scale-up synthesis of compound **5**.

Due to the excellent research and development prospects of **5**, pilot-scale synthesis of **5** must be solved. If we use the previously reported synthetic route of **1** to synthesize **5**,²² a total yield of less than 1% would be obtained, and finally only mg-level product could be acquired. Besides the low selectivity of demethylation for intermediate **11** and low condensation efficiency of intermediate **7**, the separation of the enantiomers by preparative HPLC in the synthetic route of **1** also hampered to provide Kg-level product for further studies.

Herein, we reported an optimized route for the preparation of **5** with the overall yield improved to 17% (Scheme 1). In this optimized approach, the AlCl_3/KI -mediated cleavage of the methyl-oxygen bond in compound **11** was proceeded regio-selectively and isolated as a white crystalline solid in more than 90% purity through a simple filtration at a yield of 89%

1
2
3
4 which is much higher than the previous 40%. The key intermediate **7** was successfully
5
6 obtained by Michael addition and Henry reaction of substituted salicylaldehyde **8** on
7
8 nitroalkene **9**. After optimization by trying out a variety of bases and ligands, we found that
9
10 the use of L-(-)-piperidine-2-carboxylic acid or morpholine instead of DABCO as catalyst
11
12 gave the corresponding coupling product **7** more effectively in 75% yield which is much
13
14 better than the ~20% of the original process, and the crude product was used in the next step
15
16 without further purification. Reduction of **7** with NaBH₄ gave a mixture of *cis/trans*
17
18 dihydrochroman derivatives **6**, then configuration transformation of the *cis* isomers to the
19
20 *trans* isomers could occur in one-pot upon exposure to DIPEA in ethanol at room
21
22 temperature. Subsequently, reduction of the nitro group of *trans* isomers of **6** to an amine
23
24 proceeded with Zn and HCl in ethanol formed enantiomers of **5** smoothly. The two-step
25
26 reaction of NaBH₄ reduction and configuration transformation was completed in one step,
27
28 and the yield was 80%, which is significantly higher than the yield of less than 50% of the
29
30 original process.
31
32
33
34
35
36
37
38
39

40 The absolute configuration of **5** was determined to be (2*S*, 3*R*) by X-ray crystallographic
41
42 analysis. Having established a scale-up method to access *trans*-enantiomers **5**, we explored
43
44 several commercially available resolving agents to afford optically pure **5**. Through a rational
45
46 screening approach, (+)-Dibenzoyl-D-tartaric acid was selected as the resolving agent to
47
48 achieve the desired salts of **5** with ee value 99% and yield 37% after re-crystallization twice
49
50 from acetone/H₂O, which is much more convenient than the separation of enantiomers **5** by
51
52 chiral preparative HPLC in the original process. The total yield of the entire synthetic route
53
54 was increased from less than 1% to nearly 20% in the present study, and the operation
55
56
57
58
59
60

1
2
3
4 process has been greatly simplified. Through the Kg-level production procedure above
5
6 described, we have finished nearly 10 Kg product of **5** for ongoing preclinical studies.
7
8
9

10 11 **Inhibitor **5** binds DPP-4 with fast-associating and slow-dissociating profiles**

12
13
14 Compound **5** inhibited DPP-4 activity (Table S1) reversibly and showed a clear
15
16 characteristic of fast-associating and slow-dissociating inhibition (Table 1, Fig. 2a) with the
17
18 association rate constant (k_{on}) of $1.29 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$ and dissociation rate constant (k_{off}) of
19
20 $2.29 \times 10^{-3} \text{ s}^{-1}$. Compared with **1**, the association rate constant of **5** has increased by an order of
21
22 magnitude, while the dissociation rate constant decreased slightly. Hence, **5** has a smaller K_{D}
23
24 value than **1** (1.77×10^{-10} vs $2.38 \times 10^{-9} \text{ M}$). Moreover, compound **5** binds tighter than sitagliptin
25
26 ($K_{\text{D}} = 1.27 \times 10^{-8} \text{ M}$) and omarigliptin ($K_{\text{D}} = 2.75 \times 10^{-9} \text{ M}$) to the DPP-4 enzyme (Fig. S3). The
27
28 kinetics studies also demonstrated that compound **5** binds faster than both sitagliptin and
29
30 omarigliptin to the DPP-4 enzyme. The fast on-rate of **5** could increase its rebinding to
31
32 DPP-4, which is the situation where a ligand, after dissociated from the target, binds to the
33
34 target again rather than diffuses away from it.^{24, 32, 33} Rebinding of **5** could lead to a high local
35
36 drug concentration, which is beneficial to decrease the rate of decline of target occupancy,
37
38 thus resulting in a prolonged *in vivo* duration of action.^{24, 34} More importantly, the
39
40 dissociation of bound **5** from the DPP-4 enzyme was 6.5 times slower than sitagliptin,
41
42 demonstrating that the residence time ($1/k_{\text{off}}$) of **5** was 6.5 times longer than sitagliptin.
43
44 Drug-target residence time has been reported to be a favorable parameter for long-lasting *in*
45
46 *vivo* efficacy.^{25, 35} Collectively, the fast on-rate and slow off-rate kinetics profiles of **5** would
47
48 ensure an extended overall duration of target occupancy, thus leading to a long-acting DPP-4
49
50
51
52
53
54
55
56
57
58
59
60

inhibition efficacy *in vivo* (Fig. S4).

Table 1. Binding kinetics profiles of 5

Compound	k_{on} ($\text{M}^{-1}\text{s}^{-1}$)	k_{off} (s^{-1})	K_{D} (M)
1	$1.44 \pm 0.022 \times 10^6$	$3.43 \pm 0.045 \times 10^{-3}$	2.38×10^{-9}
5	$1.29 \pm 0.055 \times 10^7$	$2.29 \pm 0.078 \times 10^{-3}$	1.77×10^{-10}
sitagliptin	$1.18 \pm 0.010 \times 10^6$	$1.49 \pm 0.012 \times 10^{-2}$	1.27×10^{-8}
omarigliptin	$2.58 \pm 0.022 \times 10^5$	$7.09 \pm 0.12 \times 10^{-4}$	2.75×10^{-9}

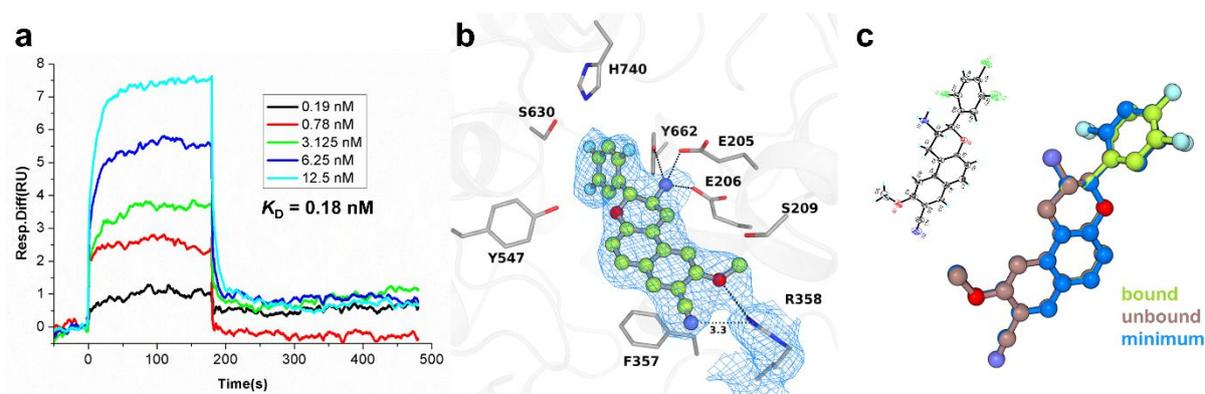


Fig. 2 (a) Kinetics analysis of the binding of **5** to DPP-4. (b) Crystal structure of DPP-4 in complex with candidate **5**. $2Fo-Fc$ electron density is contoured at 1σ . The electron densities of both **5** and Arg358 are clear, and their position and pose are unambiguous. (c) Single crystal X-ray structure of **5**, and comparison of the bound, unbound and global minimum conformations of **5**. The three conformations are superimposed well with each other.

Inhibitor **5** binds DPP-4 with Arg358 in a closed-up conformation

In order to verify whether **5** makes close dipole-dipole interactions with residue Arg358 in a closed-up conformation as expected, we solved the X-ray crystal structure of **5** in complex with DPP-4 (Table S2). The complex crystal structure (Fig. 2b) shows us that **5** binds at the catalytic pocket of the DPP-4 enzyme with its amino group on the dihydropyran

ring positioned appropriately to generate charge-reinforced hydrogen-bonding interactions with the side chains of Glu205, Glu206 and Tyr662 of the recognition site for the peptide substrate N-terminus. The 2,4,5-trifluorobenzene warhead penetrates into the hydrophobic S1 subsite adjacent to the catalytic triad (Ser630, Asp708, and His740). The naphthalene fragment of **5** contacts Phe357 with favorable face-to-face π - π stacking interactions. Our previous work had found that shape matching and electrostatic complementarity are two crucial factors for DPP-4 inhibitors to accommodate the widely opened S2 extended subsite that covered with relatively strong positive electrostatic potential (Fig. S5).²² For inhibitor **5**, at the S2 extended subsite, the -OCH₃ at R¹ could insert into the concave surface formed by residues Glu206, Val207 Ser209 and Arg358 to make favorable van der Waals (vdW) interactions. Accompanied by -OCH₃ at R¹, the electron-withdrawing -CN at R² attracts the flexible Arg358 and forms charge-reinforced hydrogen bond with it just as expected. Moreover, the electron densities of Arg358 are clear, and its position and pose in a closed-up pattern are unambiguous. Collectively, those binding characteristics made **5** a very tight-binding DPP-4 inhibitor.

By superimposing the modeled complex structure of **5** in DPP-4 with the X-ray crystal structure, an RMSD value of 0.246 Å was obtained for the pocket-lining residues with both Arg358 and **5** aligned well with each other (Fig. S6), suggesting that the FMO-QM based drug design strategy is reasonable.

Negligible conformational reorganization between unbound and bound state of **5**

The introduction of a linear and rigid -CN at R² of our scaffold not only provides a

1
2
3
4 functional group that generates relatively strong electrostatic interactions with Arg358 but
5
6 also makes **5** more rigid by reducing one rotatable bond compared with our -OCH₃
7
8 substituted lead compound **1**. In solution, **5** preferentially assumes the same liner-shaped
9
10 conformation (Fig. 2c, Fig. S7, Table S3) as that required for DPP-4 complexation. The
11
12 RMSD value between the unbound and bound state of **5** is 0.01 Å, indicating that **5** incurred
13
14 almost no conformational reorganization when going from the solution unbound state to the
15
16 DPP-4 bound state. Thus, little reorganization free energy penalty is needed to be
17
18 counterbalanced by the strong intermolecular interactions of the ligand and receptor, resulting
19
20 in a subnanomolar binding affinity of **5**. Furthermore, both the unbound and bound
21
22 conformations of **5** are very similar to its global minimum conformation (RMSD < 0.16 Å,
23
24 Fig. 2c, Table S4), as are their conformational energies (Δ SPE < 0.32 kcal/mol, Table S5).
25
26 The negligible conformational strain energy of **5** in its bioactive conformation could greatly
27
28 contribute to its tight and stable binding to DPP-4.
29
30
31
32
33
34
35
36
37
38
39

40 **Inhibitor 5 has a long half-life across preclinical species**

41
42
43 Compound **5** has a long half-life (> 24h) in both mouse and rat (Table 3). In ICR mice, **5**
44
45 achieved the maximal plasma concentration slowly (T_{\max} = 10.20 h), displayed a large
46
47 exposure (AUCs), and had a very long terminal half-life of 25.31 h for PO dosing. In SD rats,
48
49 **5** had a larger exposure (15377.72 vs 3415.00 µg/L*h), wider distribution (2.07 vs 1.66 L/kg)
50
51 and slower plasma clearance rate (0.066 vs 0.30 L/h/kg) than **1**, leading to a longer terminal
52
53 half-life ($T_{1/2}$) for iv dosing (11.99 vs 3.87 h). While for the PO administration, $T_{1/2}$ was not
54
55 observed within the experimental time period of 24 h, which indicated that **5** can maintain a
56
57
58
59
60

high plasma concentration for a long time (> 24 h) *in vivo*. In addition, the terminal half-life of **5** is much longer than sitagliptin (1.7 h) and omarigliptin (11 h) in rats with the same doses for iv and po administration.^{36, 37} The sustained high plasma concentration of **5** would undoubtedly contribute to a long duration of action as an inhibitor of plasma DPP-4 *in vivo*. The oral bioavailability of **5** is very high (98.54%) in rats, indicating that **5** has good gastrointestinal permeability.

Table 3. Preclinical pharmacokinetic profiles of 5.

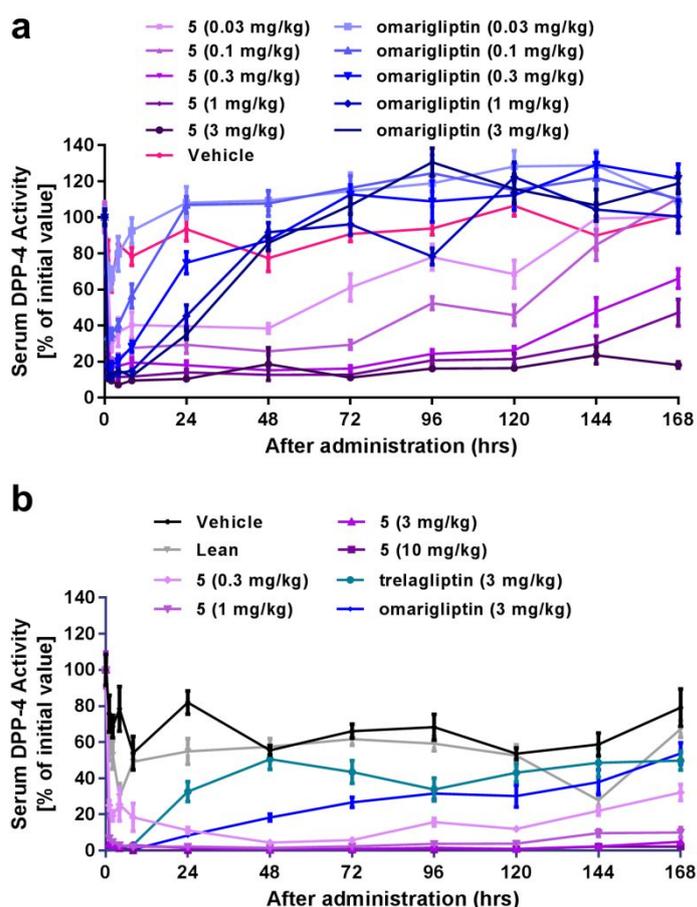
	Group		
	Mouse ^a	Rat ^b	
Administration	po	po	iv
Dose (mg/kg)	2	2	1
$T_{1/2}$ (h)	25.31	NC ^c	11.99 ± 1.25
T_{max} (h)	10.20	6.67 ± 1.15	NC
C_0/C_{max} (µg/L)	1773.37	1387.47 ± 145.99	2007.65 ± 217.16
AUC _{0-t} (µg/L*h)	82006.97	23162.26 ± 3041.78	11752.88 ± 962.01
AUC _{0-∞} (µg/L*h)	85621.45	NC	15377.72 ± 1776.19
Cl (L/h/kg)	NC	NC	0.066 ± 0.007
MRT _{0-∞} (h)	40.68	NC	15.87 ± 1.96
Vd _{ss} (L/kg)	NC	NC	2.07 ± 0.18

^a Values are mean (n = 4). ^b Values are mean ± SD (n = 3). ^c Not observed within the experimental time period of 24 h. NC: Not calculated.

Pharmacodynamic profiling in ICR mice, type 2 diabetic *ob/ob* mice

On the basis of the superior properties of **5** including potent inhibitory potency *in vitro*, long residence time and favorable pharmacokinetic characteristics, it was advanced into further *in vivo* efficacy study. First, the plasma DPP-4 inhibition study of **5** was primarily evaluated in ICR mice. As shown in Figure 3a, single dose oral treatment of **5** resulted in a

1
2
3
4 time and dose-dependent inhibition of plasma DPP-4 activity in ICR mice, and **5** (3 mg/kg)
5
6 displayed sustained up to 80% inhibitory effect of DPP-4 activity in 168 h (7 days) after
7
8 single dose administration, which is superior to the performance of the long-acting control
9
10 omarigliptin (3 mg/kg). Those results inspired us to conduct another experiment to further
11
12 validate whether **5** can have a sustained bioactivity against DPP-4 in type 2 diabetic *ob/ob*
13
14 mice (Fig. S8). As a result, **5** inhibited the DPP-4 activity in a dose-dependent manner (Fig.
15
16 3b), and >80% inhibitory effect of DPP-4 activity was observed through 168 h after single
17
18 dose administration, which is better than trelagliptin and omarigliptin.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55



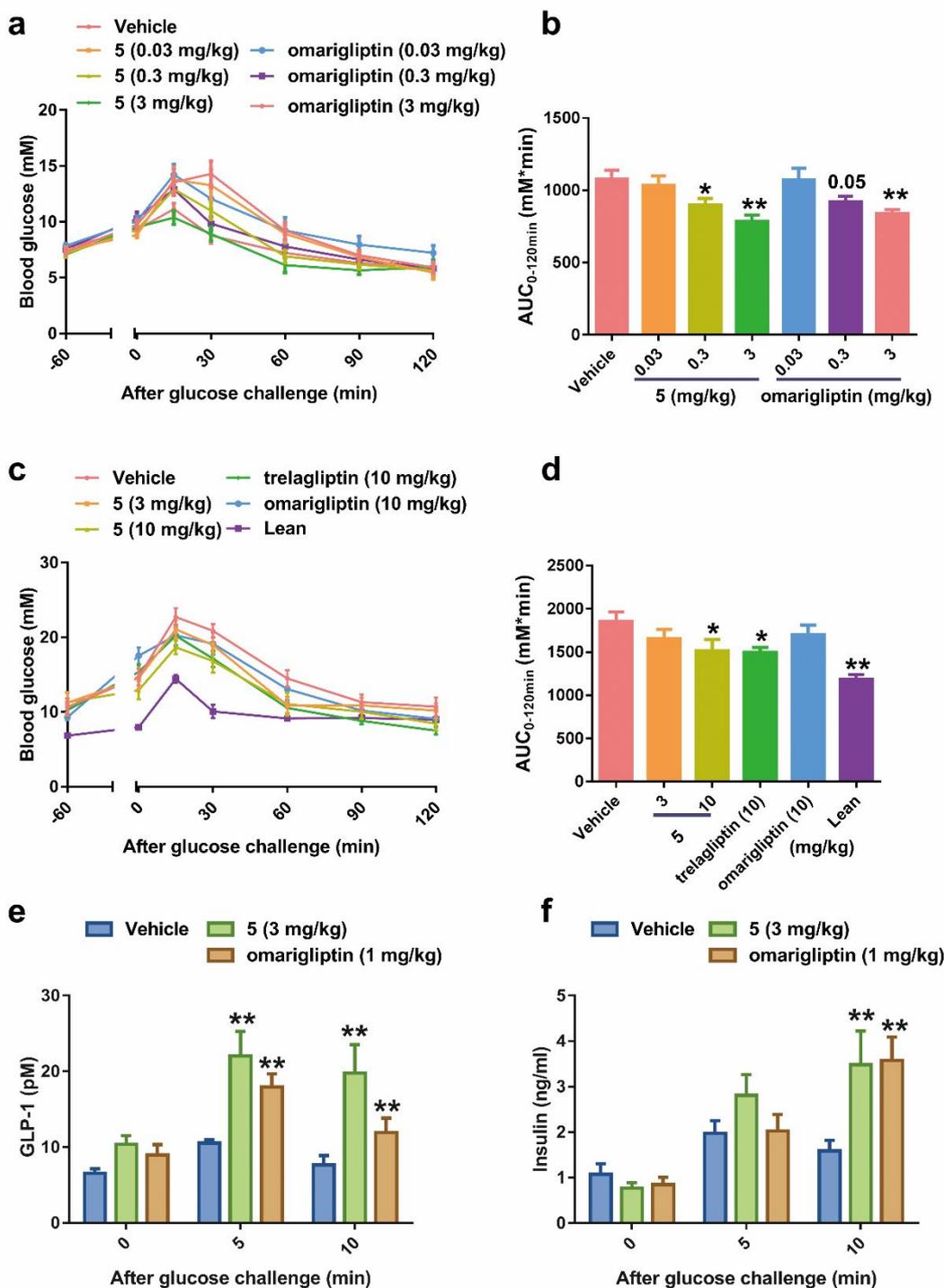
56 **Fig. 3** Inhibition of DPP-4 activity in serum or plasma obtained from the normal and diabetic
57 animal. (a) Inhibition of DPP-4 activity in serum obtained from ICR mice after a single oral
58
59
60

1
2
3
4 administration of **5** and omarigliptin, respectively (n = 10, male). Data are presented as the
5
6 mean \pm SE. (b) Inhibition of DPP-4 activity in serum obtained from type 2 diabetic *ob/ob*
7
8 mice after a single oral administration of **5**, trelagliptin and omarigliptin, respectively (n = 10,
9
10 5 male and 5 female). Data are presented as the mean \pm SE.
11
12
13
14
15
16

17 **Effect of acute administration of 5 on serum glucose, active GLP-1 and insulin levels**

18
19 Inhibition of DPP-4 augments the level of active GLP-1 and insulin by inhibiting the
20
21 degradation of active GLP-1, and it returns glucose homeostasis toward physiological control
22
23 levels.³⁸ Considering the favorable *in vivo* DPP-4 inhibition activity of **5**, it was further
24
25 advanced to evaluate its ability to increase active GLP-1 levels, insulin levels and improve
26
27 the glucose clearance capacity. In ICR mice, **5** administered 60 min before a glucose
28
29 challenge markedly reduced blood glucose AUC_{0-120 min} by 16.4% and 27.2% at the dose of
30
31 0.3 and 3 mg/kg (P < 0.05), respectively, which were similar with omarigliptin (14.6% and
32
33 22.1%) (Fig. 4a-4b). Oral glucose tolerance test (OGTT) in *db/db* mice showed that
34
35 single-dose oral treatment with **5** (10 mg/kg) significantly improved the glucose tolerance
36
37 capacity by decreasing the AUC_{0-120 min} of 18.3% compared with vehicle group (P < 0.05), and
38
39 the pharmacodynamic effect was equipotent to trelagliptin (Δ AUC_{0-120 min} = 19.3%) but
40
41 superior to omarigliptin (Δ AUC_{0-120 min} = 8.3%) (Fig. 4c-4d, Fig. S9). We also evaluated the
42
43 GLP-1 and insulin secretion effect of **5**. Upon treatment of ICR mice with **5** at 3 mg/kg,
44
45 active GLP-1 level significantly increased by 2.0 and 2.6-fold compared with the vehicle
46
47 control (P < 0.01) at 5 min and 10 min after the glucose challenge, which were better than
48
49 omarigliptin(1.7 and 1.6-fol) (Fig. 4e). Likewise, the insulin levels were also increased by
50
51
52
53
54
55
56
57
58
59
60

administration of **5**, which is comparable to the efficacy of omarigliptin (Fig. 4f). This demonstrated that a single oral dose of **5** inhibited serum DPP-4 activity, which increases active GLP-1 and insulin levels and improves glucose homeostasis in normal mice and genetic-type 2 diabetic mice.



1
2
3
4 **Fig. 4** Effects of a single oral administration of **5** on plasma glucose, active GLP-1, and
5
6 insulin levels. (a-b) Blood glucose changes in OGTT and AUC of blood glucose between 0
7
8 and 120 min in ICR mice (n = 10, male). Compound **5**, trelagliptin, omarigliptin or vehicle
9
10 was orally administered to 8h-fasted ICR mice 60 min prior to oral glucose challenge (3
11
12 g/kg). Data are presented as the mean \pm SE. (c-d) Blood glucose changes in OGTT and AUC
13
14 of blood glucose between 0 and 120 min in *db/db* mice (n = 10, 5 male and 5 female).
15
16 Compound **5**, trelagliptin, omarigliptin or vehicle was orally administered to 6h-fasted *db/db*
17
18 mice 60 min prior to an oral glucose challenge (1.5 g/kg). Data are presented as the mean \pm
19
20 SE. (e-f) Plasma active GLP-1 and insulin levels following OGTT in ICR mice (n = 10,
21
22 male). Compound **5**, trelagliptin, omarigliptin or vehicle was orally administered to 8h-fasted
23
24 ICR mice 60 min prior to an oral glucose load (3 g/kg). Data are presented as the mean \pm SE.
25
26 (*) P < 0.05 vs vehicle; (**) P < 0.01 vs vehicle. (b.d) One-Way ANOVA with Fisher's LSD
27
28 test for multi-comparisons. (e.f) Two-Way ANOVA with Fisher's LSD test for
29
30 multi-comparisons.
31
32
33
34
35
36
37
38
39
40
41
42

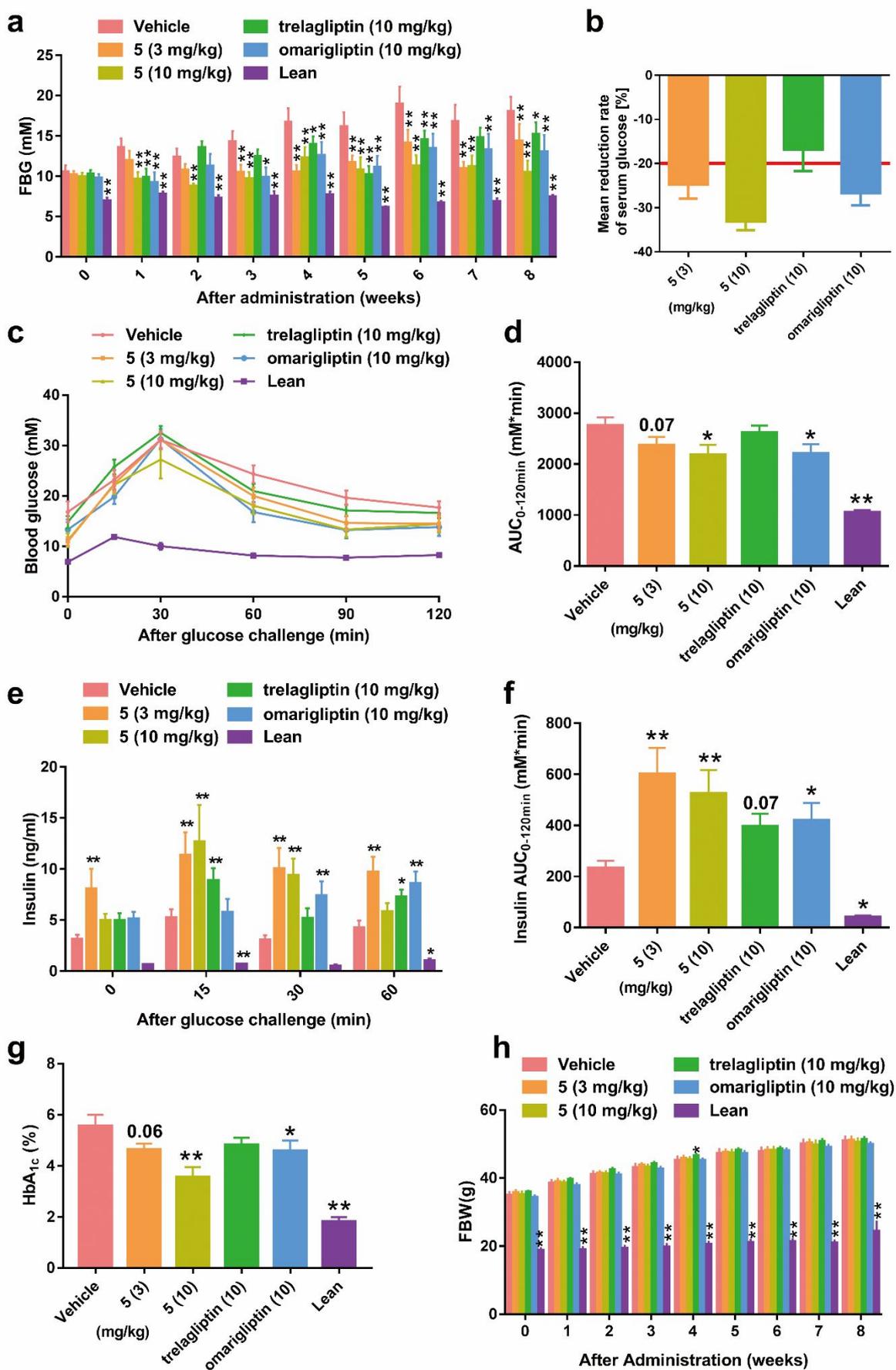
43 **Chronic antidiabetic effects of 5 on type 2 diabetic *db/db* mice**

44
45 Encouraged by the results from the acute glucose-lowering activity, long-term
46
47 antidiabetic effects of **5** once weekly were evaluated in genetic-type 2 diabetes animal models
48
49 (3 and 10 mg/kg, q.w) for 8-weeks. Administration of **5** with the dose of 3 and 10
50
51 mg/kg/week caused significant reductions in FBG levels (Fig. 5a-b), and the average
52
53 reduction during the entire treatment period was 24.7% and 33.1%, respectively, which were
54
55 better than trelagliptin (16.8%) and omarigliptin (26.6%). Chronic administration of **5** (10
56
57
58
59
60

1
2
3
4 mg/kg/week) in *db/db* mice also significantly improved the glucose tolerance of the diabetic
5
6 mice ($P < 0.05$), which is comparable to omarigliptin and better than trelagliptin (Fig. 5c-d).
7
8 Remarkably, an glucose stimulated insulin secretion test revealed that treatment with **5**
9
10 significantly increased insulin level and raised by 2.6 and 2.2-fold compared to the vehicle
11
12 control ($P < 0.01$) in $AUC_{0-120 \text{ min}}$, which are much larger than those of trelagliptin and
13
14 omarigliptin at the same dose of 10 mg/kg/week (1.7 and 1.8-fold vs the vehicle control, $P <$
15
16 0.05) (Fig. 5e-f). Notably, the 10 mg/kg/week **5** performed the best (HbA1c% = 3.56%, $P <$
17
18 0.01) when compared with the long-acting controls trelagliptin and omarigliptin (Fig. 5g).
19
20 Collectively, the long-term antidiabetic effect of **5** for once-weekly treatment is superior to
21
22 both trelagliptin and omarigliptin in *db/db* mice. Since trelagliptin and omarigliptin were
23
24 approved for antidiabetic treatment with a once-weekly dosing regimen, our candidate **5**, with
25
26 longer and better antidiabetic efficacies in preclinical species than both trelagliptin and
27
28 omarigliptin, is likely to be amenable for once-weekly dosing in human.
29
30
31
32
33
34
35
36
37

38 GLP-1 and its analogs have been demonstrated to reduce the body weight in both
39
40 diabetic animal models and clinical studies. Inhibition of DPP-4 activity prevented
41
42 degradation and enhanced the biological activity of active GLP-1.³⁹ However, the fasting
43
44 body weight and postprandial body weight after chronic administration of **5** increased
45
46 gradually as those of vehicle, omarigliptin and trelagliptin, demonstrating that **5** (both 3 and
47
48 10 mg/kg/week) is well tolerated in *db/db* mice, and do not cause abnormally weight gain.
49
50 These results are similar to the reports of several other studies of DPP-4 inhibitors. The
51
52 possible reasons might be due to the less effect of the enhanced endogenous active GLP-1
53
54 caused by DPP-4 inhibition rather than that of the injection of exogenous GLP-1 or its
55
56
57
58
59
60

1
2
3
4 analogs.³⁹ In addition, all animals tolerated exposure to therapeutic concentrations of
5
6 compound **5** without any signs of distress or discomfort.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



1
2
3
4 **Fig. 5** Long-term antidiabetic efficacies of **5** in diabetic *db/db* mice with 3 and 10
5 mg/kg/week for 8 weeks (n = 10, 5 male and 5 female). (a) Fasting blood glucose (FBG)
6 changes, (b) Mean reduction rate of FBG, (c) Blood glucose changes in OGTT, (d) AUC of
7 blood glucose between 0 and 120 min (the areas under the curve are indicators of glucose
8 clearance), (e) Plasma insulin levels following OGTT, (f) AUC of insulin between 0 and 60
9 min (the areas under the curve are indicators of insulin secretion), (g) Plasma HbA1c levels,
10 and (h) Fasting body weight (FBW). OGTT was conducted after 7 weeks treatment of **5** after
11 an oral glucose load (1.5 g/kg). Plasma insulin levels after an oral glucose load (1.5 g/kg)
12 were performed after 8 weeks treatment of **5**. Data are presented as the mean \pm SE. (*) P <
13 0.05 vs vehicle; (**) P < 0.01 vs vehicle. (a,e,h) Two-Way ANOVA with Fisher's LSD test
14 for multi-comparisons. (d,f,g) One-Way ANOVA with Fisher's LSD test for
15 multi-comparisons.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

38 **Preclinical safety studies**

39
40 Blockage of the human potassium channel hERG (human Ether-à-go-go-Related Gene)
41 can prolong the cardiac action potential and the QT-interval, which is a major concern in drug
42 discovery and development. The selectivity of **5** against hERG was tested using Patch-clamp
43 experiments, and the results showed that the IC₅₀ value of **5** on hERG is larger than 40.00
44 μ M.
45
46
47
48
49
50
51
52

53 We also conducted preclinical safety studies of **5** in ICR mice. For a single dose oral
54 toxicity study, **5** was well tolerated at 300 mg/kg with no obvious toxicity. For 14
55 consecutive days of oral toxicity study, **5** was well tolerated at a dose of 30 mg/kg once daily
56
57
58
59
60

1
2
3
4 based on clinical signs, but no histopathology was performed. Body weight and organ index,
5
6 blood parameter examination and serum biochemical metabolites of the ICR mice were
7
8 comparable to those of the control group (Table S6-S9).
9
10

11 12 13 14 15 **Discussion and Conclusions**

16
17
18 Our previously discovered inhibitor **1** unexpectedly induced Arg358 to form a closed-up
19
20 conformation, which may be one crucial structural determinant for **1**'s slow binding kinetics
21
22 profile. With the attempt to optimize both the binding kinetics and pharmacokinetic
23
24 properties of **1** to get a super long-acting antidiabetic candidate for once-weekly dosing, we
25
26 intended to lock Arg358 in a closed-up conformation in the following structural optimization.
27
28 The FMO-QM method helped us quantitatively evaluate the specific contribution made by
29
30 the two -OCH₃ groups in **1** that attract Arg358, and further guided us to perform structural
31
32 optimization at R². By elaborately introducing a linear 2.4 Å -CN group, which is an
33
34 electronegative H-bond acceptor, to mimic the role of the -OCH₃ group at R² cooperating
35
36 with the other -OCH₃ at R¹ in **1** to attract Arg358, **5** displayed a slightly stronger electronic,
37
38 charge transfer and dispersion interactions with Arg358 than **1**. X-ray crystal structure of **5** in
39
40 complex with DPP-4 disclosed that **5** is expectedly inducing Arg358 into the same closed-up
41
42 conformation as that in the complex crystal structure of **1**. Subsequent SPR assay
43
44 demonstrated that **5** binds faster and tighter to DPP-4 than **1** and sitagliptin, while the
45
46 disassociating rate is slightly slower than **1** but an order of magnitude slower than sitagliptin.
47
48 Pharmacokinetic studies provided that **5** has a larger exposure, wider distribution and slower
49
50 clearance rate than **1**, thus leading to a longer half-life than **1**. Those distinctive binding
51
52
53
54
55
56
57
58
59
60

1
2
3
4 kinetics and pharmacokinetic characteristics made **5** a potent and orally available DPP-4
5
6 inhibitor with longer and better antidiabetic efficacy than **1** and the once-weekly controls
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

kinetics and pharmacokinetic characteristics made **5** a potent and orally available DPP-4 inhibitor with longer and better antidiabetic efficacy than **1** and the once-weekly controls trelagliptin and omarigliptin in preclinical species, suggesting that compound **5** holds promise for an efficacious and safe clinical candidate for the once-weekly treatment of T2DM to improve ease of use and compliance.

Both the fast-associating and slow-dissociating binding kinetics profiles and the pharmacokinetic profiles of **5** contribute to its long-acting antidiabetic efficacy *in vivo*. The fast on-rate of **5** can increase its rebinding to DPP-4. Rebinding is helpful to maintain a high degree of target occupancy. The slow off-rate of **5** can ensure a long drug-target residence time, which is beneficial for a long duration of action *in vivo*. The long terminal half-life is favorable for sustaining a high plasma concentration of **5** to inhibit the soluble DPP-4 enzyme in the blood. With a large volume of distribution, **5** may also inhibit the membrane-bound DPP-4 in the tissues, such as liver and small intestine. This may also be conducive to the long duration of action. Therefore, both the special binding kinetic profiles and pharmacokinetic profiles are supposed to be main drivers of the long-lasting antidiabetic efficacy of **5** as an inhibitor of DPP-4 *in vivo*.

Arg358 is one of the most important residues of the S2 extended subsite, which is flexible in apo and holo structures of the DPP-4 enzyme,⁴⁰ and its orientation influences the shape and physicochemical properties of the S2 extensive subsite of DPP-4.⁴¹ Although many DPP-4 inhibitors have been discovered, as far as we know, there is no specific structure-based drug design work conducted against Arg358 with the aim of fixing its flexible sidechain in a closed-up conformation. In our case, we successfully locked Arg358 in a

1
2
3
4 closed-up conformation to accommodate the substitutions at R¹ and R² of our scaffold by
5
6 using the FMO method. And this induced conformational change of Arg358 is supposed to be
7
8 one structural determinant that contributes to the slow dissociation kinetics of the inhibitors.
9
10 The FMO-QM method can be served as a useful tool in guiding structural optimization where
11
12 some key residues are preferentially targeted in a ligand-protein complex.
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Experimental Section

1. Computational Methods

The FMO-QM method was developed by Kitaura and co-workers, which can accurately quantify the binding interactions of ligand-residue pairs.²⁸⁻³¹ It has been successfully applied in many biological systems to study ligand-protein interactions, in which the *ab initio* FMO method clearly outperformed the molecular mechanics (MM) based scoring functions in predicting binding energies.⁴²⁻⁴⁵ Compared with MM based methods, the major advantages of FMO method is its consideration of dynamic polarization and charge transfer of the ligand-protein system when computing electrostatic interaction.

Structure Preparation. Molecular Operating Environment (MOE, version 2014.09) program was used to prepare the complex structure of the ligand and DPP-4 for the FMO calculation. All water molecules were removed. The missing hydrogen atoms of the crystal structure of **1** in complex with DPP-4 (PDB: 5J3J, chain B) were added and the protonation state of the acidic and basic amino acid residues was predicted using the Protonate 3D module. Default settings including pH of 7.0, a cutoff of 15 Å for the generalized Born/volume integral methodology and solvent dielectric constant of 80 were applied. Then, a structural minimization was conducted to the complex with each atom restrained to deviate up to 0.5 Å from its original position in the crystal structure by using the semi-empirical Amber10:EHT forced field that implemented in the MOE program. The refined complex structure was then used for the subsequent FMO calculation.

Compounds **2-5** were close analogs to **1**, so they were likely to exhibit a binding pattern similar to **1**. Therefore, we generated the binding conformations for compounds **2-5** in

1
2
3
4 complex with DPP-4 in two steps: 1) superimpose the inhibitor to the binding pose of **1** in the
5
6 crystal structure using SHAFTS,^{46, 47} which is an in-house software for 3D small molecular
7
8 similarity calculation and structural alignment based on both feature and shape matches; 2)
9
10 prepare and minimize the complex structure using the procedures described above.
11
12

13
14 *FMO calculation.* In general, the FMO calculation consists of four steps: 1) Fragment a
15
16 large molecular system into a number of small monomer fragments; 2) Calculate *ab initio*
17
18 molecular orbital on each monomer (fragment) in the Coulomb field exerted by all remaining
19
20 monomers repeatedly until all monomer electron densities become self-consistent, whereby
21
22 intra-fragment charge transfer is taken into account; 3) Compute self-consistent field (SCF)
23
24 for each dimer (fragment pair), through which inter-fragment charge transfer and other
25
26 quantum effects are accounted for; 4) Calculate inter-fragment interaction energy and total
27
28 energy of the whole system. The detailed principle and methodology of FMO can be found in
29
30 the literature.²⁸⁻³¹
31
32
33
34
35
36

37
38 In this study, the FMO input files were prepared in Facio (version 19.1.5) to include
39
40 residues around 4.5 Å of the ligand and cut them off into one-residue-per-fragment at C α of
41
42 each residue. The ligand was also treated as one fragment. All FMO calculations were
43
44 performed using GAMESS (version Dec, 2014),^{48, 49} which is a general *ab initio* quantum
45
46 chemistry package, at the correlated second-order Møller-Plesset (MP2) level⁵⁰ with the
47
48 6-31G* basis set. Because all the FMO calculations in the present work were performed in
49
50 vacuum condition, we treated the amino group in each DPP-4 inhibitor as uncharged NH₂ to
51
52 avoid overestimation of the charge-charge electrostatic interactions.
53
54
55
56
57

58
59 After FMO calculation, a list of pair interaction energies (PIEs) between fragments was
60

1
2
3
4 obtained. The PIE between any two fragments i and j ($\Delta E_{ij}^{\text{int}}$) is composed of four energy
5
6 terms (eq 1): electrostatics ($\Delta E_{ij}^{\text{es}}$), charge-transfer ($\Delta E_{ij}^{\text{ct}}$), dispersion ($\Delta E_{ij}^{\text{di}}$) and
7
8 exchange-repulsion ($\Delta E_{ij}^{\text{ex}}$):
9

$$\Delta E_{ij}^{\text{int}} = \Delta E_{ij}^{\text{es}} + \Delta E_{ij}^{\text{ct}} + \Delta E_{ij}^{\text{di}} + \Delta E_{ij}^{\text{ex}} \quad (1)$$

10
11
12
13
14 In our study, fragment i is the ligand, and j is any one of the pocket-lining residues. It is
15
16 worth noting that the Δ symbols refer to the differences between the total QM energies of a
17
18 fragment pair ij and two individual fragments i and j , both of which are computed at the
19
20 bound state of the protein-ligand complex. Thus, the sum of the PIEs of the system represents
21
22 the binding stability of the protein-ligand complex, which is correlated with but not the same
23
24 as the binding free energy.^{49, 51}
25
26
27
28
29
30
31

32 2. Animals study

33
34
35 Male ICR mice (7-8 weeks) and male SD rats (7-8 weeks) were purchased from
36
37 Shanghai SLAC Laboratory Animal Co., Ltd. B6.V-Lep^{ob/+} (ob/+) mice and
38
39 BKS.Cg-m^{+/+}Lepr^{db/J}(db/+) were imported from Jackson Laboratory (Bar Harbor, ME,
40
41 USA) and bred in-house for obtaining the spontaneous diabetic *ob/ob* mice and *db/db* mice,
42
43 and littermates for control at the Shanghai Institute of Materia Medica, Chinese Academy of
44
45 Sciences. The animals were maintained under a 12-h light-dark cycle with free access to
46
47 water and food. Animal experiments were approved by the Animal Care and Use Committee,
48
49 Shanghai Institute of Materia Medica, Chinese Academy of Sciences (IACUC Number:
50
51 2015-04-LJ-10).
52
53
54
55
56
57
58
59
60

1
2
3
4 **3. In vitro bioassay, crystal structure determination and pharmacokinetic assay in SD**
5
6 **rats**
7

8
9 The *in vitro* DPP-4 inhibition study (at least three independent experiments), binding
10 kinetics study using surface plasmon resonance, the cocrystallization of DPP-4 with
11 compound **5** as well as structure determination, and the pharmacokinetic assay in SD rats
12 were all conducted using the same method of operation reported in our previous work.²²
13
14
15
16
17
18
19
20
21

22 **4. Pharmacokinetic studies in ICR mice**
23

24 Male ICR mice (n = 4 in each group) were randomly assigned to treatment groups. After
25 2 h-fasted, baseline blood was collected into a tube containing EDTA. Animals were then
26 treated orally with compound **5** (2 mg/kg). Subsequently, blood per animal was collected at 1,
27 2, 4, 6, 8, 24, 48, 72, 96 and 120 h. All samples were centrifuged at 10000 rpm for 2 min, and
28 the plasma was harvested. Aliquots of plasma samples were stored at -80 °C until analysis.
29 Compound **5** serum concentrations were determined by liquid chromatography/tandem mass
30 spectrometry (LC/MS/MS). Pharmacokinetic parameters were determined from compound **5**
31 serum concentrations by noncompartmental methods using WinNonLin® Professional
32 Version 4.1 (Pharsight Corp., Mountain View, CA).
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51 **5. Effects on DPP-4 activity in ICR mice**
52

53 Male mice were randomly assigned to treatment groups (n = 10 in each group). After 2
54 h-fasted, baseline blood was collected into a tube containing EDTA. Animals were then
55 treated orally with vehicle (0.5% sodium carboxymethyl cellulose, 10 mL/kg), compound **5**
56
57
58
59
60

1
2
3
4 (0.03, 0.1, 0.3, 1 and 3 mg/kg, 10 mL/kg) or omarigliptin (0.03, 0.1, 0.3, 1 and 3 mg/kg, 10
5
6 mL/kg). Subsequently, blood per animal was collected at 1, 2, 4, 8, 12, 24, 48, 72, 96, 120,
7
8
9 144 and 168 h. All samples were centrifuged at 10000 rpm for 2 min, and the plasma was
10
11 harvested. Aliquots of plasma samples were stored at -80 °C until analysis.

12
13
14 For the measurement of *in vivo* DPP-4 activity, a continuous fluorometric assay using
15
16 Gly-Pro-AMC as substrates was applied, which can release the fluorescent AMC group after
17
18 being cleaved by DPP-4. For the plasma samples, the reaction solution was made up of 50%
19
20 plasma samples, 125 μM Gly-Pro-AMC, and buffer (100 mM
21
22 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.5, 0.1 mg/mL BSA) in a
23
24 total reaction volume of 20 μL. For the serum samples, the reaction solution was made up of
25
26 5 μL serum samples, 10 μM Gly-Pro-AMC, and buffer (100 mM HEPES, pH 7.5, 0.1 mg/mL
27
28 BSA) in a total reaction volume of 10 μL. Liberation of AMC was monitored continuously in
29
30 a 384-well plate in 5 min at 37 °C using an excitation wavelength of 360 nm and an emission
31
32 wavelength of 460 nm. The data are reported as % activity calculated as follows: % activity =
33
34 $(V_{ts}/V_{cs}) \times 100$, where V_{ts} is the rate of reaction of treated sample and V_{cs} is the rate of
35
36 reaction of each animal's baseline level (the value obtained immediately before compound
37
38 administration).

39 40 41 42 43 44 45 46 47 48 49 50 51 **6. Effect on oral glucose tolerance test in ICR mice**

52
53 To examine the effect of compound **5** on blood glucose after an oral glucose challenge
54
55 in male ICR mice (n=10 in each group), compound **5** (0.03, 0.3, 3 mg/kg), omarigliptin (0.03,
56
57 0.3, 3 mg/kg) or vehicle (0.5% sodium carboxymethyl cellulose) was orally administered to 8
58
59
60

1
2
3
4 h-fasted ICR mice 1 h prior to the oral glucose challenge (3 g/kg). Blood glucose was
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

h-fasted ICR mice 1 h prior to the oral glucose challenge (3 g/kg). Blood glucose was estimated using a glucometer (Accu-Chek; Roche) at 60 min before the glucose load and 0, 15, 30, 60, 90 and 120 min post-glucose challenge. The area under the curve (AUC) for the glucose tolerance test was calculated using the trapezoidal method. To examine the acute effect of compound **5** (0.3, 1, 3 mg/kg) and omarigliptin (1 mg/kg) on the serum active GLP-1 level and insulin level, blood samples were collected 0, 5 and 10 min after the glucose challenge and placed in Eppendorf tubes containing the DPP-4 inhibitor valine pyrrolidide (Millipore, DPP-4-010) with a final concentration of 1% blood samples and 25 mg/mL EDTA to measure serum active GLP-1[7-36 amide] levels (Millipore, EGLP-35K) and insulin levels (Millipore, EZRMI-13BK).

7. Effect on DPP-4 activity in *ob/ob* mice

Eight-week-old *ob/ob* mice (n = 10 in each group, 5 male and 5 female) were randomly assigned to treatment groups. After 2 h-fasted, baseline blood was collected into a tube containing EDTA. Mice were then treated orally with vehicle (0.5% sodium carboxymethyl cellulose, 10 mL/kg), compound **5** (0.3, 1, 3, 1 and 10 mg/kg), omarigliptin (3 mg/kg) or trelagliptin (3 mg/kg). Subsequently, blood per animal was collected at 1, 2, 4, 8, 12, 24, 48, 72, 96, 120, 144 and 168 h. All samples were centrifuged at 10000 rpm for 2 min, and the plasma was harvested. Aliquots of plasma samples were stored at -80 °C until analysis. The measurement of *in vivo* DPP-4 activity was the same as the method with ICR mice.

8. Effect on oral glucose tolerance test in *db/db* mice

1
2
3
4 To examine the effect of compound **5** on blood glucose after an oral glucose challenge
5
6 in six-week-old *db/db* mice (n=10 in each group, 5 male and 5 female), compound **5** (3 and
7
8 10 mg/kg), omarigliptin (10 mg/kg), tregliptin (10 mg/kg) or vehicle (0.5% sodium
9
10 carboxymethyl cellulose) was orally administered to 6h-fasted *db/db* mice 60 min prior to the
11
12 oral glucose challenge (1.5 g/kg). Blood glucose was estimated using a glucometer at 60 min
13
14 before the glucose load and 0, 15, 30, 60, 90 and 120 min post-glucose challenge. The area
15
16 under the curve (AUC) for the glucose tolerance test was calculated using the trapezoidal
17
18 method.
19
20
21
22
23
24
25
26

27 **9. Long-term antidiabetic effects in *db/db* mice**

28
29
30 Six-week-old *db/db* mice were divided into 5 groups (n=10 in each group, 5 male and 5
31
32 female) based on non-fasting and 6h-fasting blood glucose, serum insulin levels, postprandial
33
34 body weight (PBW, non-fasting body weight) and 6-fasting body weight (FBW). Lean
35
36 littermates were used as the lean control. Compound **5** (3 and 10 mg/kg), omarigliptin (10
37
38 mg/kg), trelagliptin (10 mg/kg) or vehicle (0.5% sodium carboxymethyl cellulose) was orally
39
40 administered once weekly for 8 weeks. Non-fasting glucose and fasting blood glucose,
41
42 postprandial body weight and 6-fasting body weight were determined at 7-d intervals. After 7
43
44 weeks of treatment, the 6h-fasted animal was challenged by 1.5 g/kg glucose. Blood glucose
45
46 was estimated using a glucometer at 0, 15, 30, 60, 90 and 120 min post-glucose challenge.
47
48 After 8 weeks of treatment, the 6h-fasted animal was challenged by 1.5 g/kg glucose. Blood
49
50 samples were collected at 0, 15, 30 and 60 min post-glucose challenge to test plasma insulin
51
52 levels. After 8 weeks of treatment, blood samples were collected after 6h of fasting for HbA1c
53
54
55
56
57
58
59
60

1
2
3
4 level measurement on the 67th day (Xin Jiang Kang Cheng, F1105). Detailed dosing regimen
5
6 was provided in the supplemental material (Fig. S11).
7
8
9

10 **10. Statistical analysis**

11
12
13
14 The results are presented as the means \pm SEM. Differences between the groups were
15
16 analyzed using one-way ANOVA or two-way ANOVA by GraphPad Prism version 7.00 for
17
18 Windows (GraphPad Software, La Jolla California USA). $P < 0.05$ was regarded as
19
20 statistically significant.
21
22
23
24
25
26

27 **11. Safety Statement**

28
29 No unexpected or unusually high safety hazards were encountered.
30
31
32
33
34

35 **12. Chemistry reagents and methods**

36
37 All commercial chemicals and solvents are reagent grade and were used without further
38
39 purification unless otherwise noted. ^1H NMR spectra were recorded on a Bruker AV-400
40
41 instrument in CDCl_3 or $\text{DMSO-}d_6$. Some ^1H NMR spectra were recorded on a Varian
42
43 InNova-500 spectrometer in $\text{DMSO-}d_6$. Data for ^1H NMR were reported as chemical shift (δ
44
45 ppm), multiplicity, coupling constant (Hz), integration. Data for ^{13}C NMR were reported as a
46
47 chemical shift. Low-resolution mass spectra were performed on a liquid
48
49 chromatography-mass spectrometer (LC-MS), using an Agilent 1200/6120 series with an
50
51 Eclipse XDB-C18 column at 25 °C. The gradient solvent system was from 80% A and 20% B
52
53 to 10% A and 90% B (A = 10 mM HCOONH_2 aqueous solution, B = CH_3CN) for 15 min
54
55
56
57
58
59
60

1
2
3
4 with a flow rate of 0.4 mL/min. High-resolution mass (HRMS, ESI) spectral data were
5
6 acquired by use of a Bruker Daltonics 7T mass spectrometer in positive ion mode and GCT
7
8 Premier HRMS instrument (EI). Purity of final compound **5** was determined with a Shimadzu
9
10 LC Solution 15C series HPLC system using C-18 column (Inertsil ODS-3 5 μ m, 4.6 mm \times
11
12 250 mm) operating at 35 $^{\circ}$ C. Elution was carried out using water as mobile phase A, and
13
14 acetonitrile as mobile phase B. Elution conditions at 0 min, phase A 20% + phase B 80%; at
15
16 25 min, phase A 20% + phase B 80%. The flow-rate of the mobile phase was 0.5 mL/min and
17
18 the injection volume of the sample was 5 μ L. The determine wavelength was 254 nm. The
19
20 purity of compound **5** was \geq 95% determined by HPLC analysis. Enantiomeric excesses (ee)
21
22 were determined by chiral HPLC analysis using Chiralcel Chiralpak AD (0.46 cm \times 25 cm).
23
24 The gradient solvent system was 80% A and 20% B for 30 min (A = Hexane, B = IPA) with a
25
26 flow rate of 0.5 mL/min. Melting point was tested using a WRS-1B digital melting point
27
28 apparatus. Flash chromatographic separations were performed using silica gel (particle size
29
30 0.040-0.045 mm). Solvents were removed by rotary evaporation under vacuum using a
31
32 standard rotovap equipped with a cold alcohol condenser. All filtrations were performed with
33
34 a vacuum unless otherwise noted.
35
36
37
38
39
40
41
42
43
44
45
46
47

48 **(2*S*,3*R*)-2-Amino-9-methoxy-3-(2,4,5-trifluorophenyl)-2,3-dihydro-1*H*-benzo[*f*]chromen**
49
50 **e-8-carbonitrile (**5**)**

51
52
53 To a stirred suspension of the racemic compounds (**6a**, **6c**) derived from last step in
54
55 ethanol (500 mL) and 6 N HCl (300 mL) was added zinc powder (157.5 g, 2.41 mol). After
56
57 being stirred for 10-24 h at RT, the mixture was filtered, and the filtrate was adjusted to pH
58
59
60

1
2
3
4 10-11 using saturated Na_2CO_3 solution and extracted with ethyl acetate (3×500 mL). The
5
6 organic solvent was washed with saturated brine and dried over anhydrous sodium sulfate
7
8 before removed by evaporation. The solution was concentrated and purified by silica gel
9
10 chromatography to get desired trans-racemic products **5** 74.5 g, 80% yield. To a solution of
11
12 the racemic product (**5**, 74 g, 192.7 mmol) in THF/ H_2O (310 mL, $V_{\text{THF}}:V_{\text{H}_2\text{O}} = 10:1$) was
13
14 added (+)-Dibenzoyl-D-tartaric acid (72.52 g, 202.4 mmol). The reaction mixture was stirred
15
16 to clarification and heated to reflux for 1 h then at RT for 1h, then the solution was stirred and
17
18 cooled down to 25 °C for 1 h before it was filtrated to give the filter cake as a white solid
19
20 while washed with cold THF. Next, to a solution of the filter cake in THF/ H_2O (220 mL,
21
22 $V_{\text{THF}}:V_{\text{H}_2\text{O}} = 10:1$) was heated to reflux for 1 h then at RT for 1 h, the solution was allowed to
23
24 come to 5 °C for 1 h and filtrated to give the qualified sample before washed with cooled
25
26 THF and dried. To a solution of the qualified sample in EtOH/ H_2O ($V_{\text{EtOH}}:V_{\text{H}_2\text{O}} = 1:10$) was
27
28 basified to a pH of 10 with a 25% aqueous solution of NaOH, the solution was allowed to
29
30 come to 30 °C for 2~3 h before filtered to give the product of **5**, the same procedure was
31
32 repeated once more to improve the enantiomeric excess of the compound. Yield **5** 27.38 g as
33
34 a white solid (37% yield, ee ≥ 99 %). Its absolute configuration was confirmed by X-ray
35
36 crystallography. Crystals suitable for X-ray crystallographic analysis were obtained from a
37
38 solution in CH_3CN by vapour diffusion. Mp: 203.7-205.1 °C . ^1H NMR (500 MHz,
39
40 $\text{DMSO}-d_6$): δ 8.40 (s, 1H), 7.75 (d, $J = 8.5$ Hz, 1H), 7.70-7.59 (m, 2H), 7.24 (s, 1H), 7.04 (d,
41
42 $J = 9.0$ Hz, 1H), 4.97 (d, $J = 7.6$ Hz, 1H), 4.04 (s, 3H), 3.45-3.36 (m, 2H), 2.83 (dd, $J_1 = 16.5$
43
44 Hz, $J_2 = 10.0$ Hz, 1H), 1.68 (s, 2H). ^{13}C NMR (100 MHz, $\text{DMSO}-d_6$): δ 157.0, 154.9, 136.9,
45
46 136.5, 128.8, 123.3, 117.8, 117.2, 113.6, 106.9, 106.7, 106.6, 106.4, 102.3, 99.5, 77.1, 56.6,
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 48.3, 31.5, 22.5. HRMS (ESI): $[M+H]^+$ 385.1158, found 385.1163. Purity >98% in HPLC
5
6 analysis.
7
8
9

10
11 ***trans*-9-Methoxy-2-nitro-3-(2,4,5-trifluorophenyl)-2,3-dihydro-1*H*-benzo[*f*]chromene-8-**
12
13 **carbonitrile (6a/6c)**
14
15

16
17 Sodium borohydride (13.75 g, 0.36 mol) was added in portions to a solution of
18 compound **7** (100.0 g, 0.24 mol) in tetrahydrofuran and methanol (1.1 L, $V_{\text{THF}}/V_{\text{CH}_3\text{OH}} = 10/1$)
19 at 15 °C. The reaction is allowed to stir at 15 °C for 2 h before quenched carefully with water
20 at the same temperature, the resulting resolution was adjusted to pH 2-3 using hydrochloric
21 acid and stirred for additional 1 h at 25 °C. The reaction was warmed to RT then isolated by
22 centrifugation and filtration to get the crude two pair of isomers. DIPEA (5 mL) was added to
23 a solution of the crude isomer racemic products in methanol (100 mL) at 60 °C under the N₂
24 atmosphere, the reaction mixture was stirred for 5-10 h. After the reaction, the excess base
25 and solvent were removed under reduced pressure to provide racemic products (**6a** and **6c**,
26 Fig. S12) as a white amorphous powder which was directly applied to next step without
27 further purification. ¹H NMR (400 MHz, DMSO-*d*₆): δ 8.48 (s, 1H), 7.85 (d, *J* = 8.9 Hz, 1H),
28 7.83-7.71 (m, 2H), 7.30 (s, 1H), 7.15 (d, *J* = 8.9 Hz, 1H), 6.01-5.80 (m, 2H), 4.05 (s, 3H),
29 3.88-3.70 (m, 2H). ¹³C NMR (100 MHz, DMSO-*d*₆): δ 157.37, 154.11, 153.52, 137.00,
30 136.04, 129.73, 129.03, 126.62, 117.86, 117.45, 116.96, 111.72, 111.00, 102.89, 102.58,
31 100.36, 81.58, 71.27, 56.92, 26.81, 25.73. LC-MS (ESI): $[M-H]^-$ *m/z* 413.08, found 413.1;
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
HRMS (EI): $[M]^+$ *m/z* 414.0827, found 414.0826.

9-Methoxy-2-nitro-3-(2,4,5-trifluorophenyl)-3H-benzo[f]chromene-8-carbonitrile (7)

A three-necked reaction kettle fitted with a water segregator, was charged with a toluene (1 L) mixture of 6-cyano-7-methoxy-2-hydroxy-1-naphthaldehyde (**8**) (100.0 g, 0.44 mol, 1.0 equiv), (E)-1,2,4-trifluoro-5-(2-nitrovinyl)benzene (**9**) (100.0 g, 0.50 mol, 1.15 equiv) and L-(-)-piperidine-2-carboxylic acid (27.97 g, 0.23 mol, 0.5 equiv). The content was heated to 120-130 °C for a period of 12-15 h under the N₂ atmosphere, the formed water was separated by a water segregator. After the reaction was complete (monitored by TLC), the slurry was then cooled to room temperature before filtration. The wet cake was washed with cold methanol to afford yellow solid 179.4 g. The solid product was collected and then diluted with CH₃OH (600 mL). The slurry was stirred at 60±5 °C for 1-2 h, after the slurry was cooled to RT, a centrifuge separate the wet cake which was next washed with methanol and dried at ambient temperature in a vacuum oven to afford yellow solid **7** 146.5 g, 75% yield.

¹H NMR (400 MHz, DMSO-*d*₆): δ 9.18 (s, 1H), 8.48 (s, 1H), 8.03 (d, *J* = 9.2 Hz, 1H), 7.92 (s, 1H), 7.76-7.70 (m, 1H), 7.62-7.55 (m, 1H), 7.10 (d, *J* = 8.8 Hz, 1H), 7.04 (s, 1H), 4.16 (s, 3H). ¹³C NMR (100 MHz, DMSO-*d*₆): δ 158.43, 155.40, 137.42, 136.91, 136.61, 135.15, 128.51, 123.90, 120.88, 117.53, 116.92, 110.70, 107.72, 103.48, 101.37, 68.05, 57.40, 55.36, 31.41, 22.52, 14.40. LC-MS (ESI): *m/z* [M+H]⁺ 413. 32, found 413.1; HRMS (EI): [M]⁺ *m/z* 412.0671, found 412.0668.

6-Cyano-7-methoxy-2-hydroxy-1-naphthaldehyde (8)

To a vigorously stirred suspension of anhydrous AlCl₃ (220.0 g, 1.64 mol) and KI (288.0 g, 1.73 mol) in dry acetonitrile (1 L) was added 6-cyano-2,7-dimethoxy-1-naphthaldehyde **11**

(100.0 g, 0.41 mol). After addition, the reaction solution was stirred at 80 °C for 4-6 h. After completion of reaction (monitored by TLC), the slurry was cooled to room temperature, the mixture was then quenched with 6 N HCl (0.7 L) and aged for additional 1 h before filtration. The desired product **8** precipitated out from the reaction mixture, the wet cake was washed with 20% acetonitrile and dried at ambient temperature in a vacuum oven to afford white solid **8** 84.0 g, 89% yield. ¹H NMR (400 MHz, DMSO-*d*₆): δ 12.14 (s, 1H), 10.68 (s, 1H), 8.58 (s, 1H), 8.35 (s, 1H), 8.03 (d, *J* = 9.2 Hz, 1H), 7.11 (d, *J* = 8.8 Hz, 1H), 3.93 (s, 3H). ¹³C NMR (100 MHz, DMSO-*d*₆): δ 192.38, 167.14, 159.33, 138.75, 136.84, 136.08, 122.11, 118.02, 116.71, 112.41, 103.01, 99.71, 56.61. LC-MS (ESI): [M-H]⁻ *m/z* 226.06, found 226.1; HRMS (EI): [M]⁺ *m/z* 227.0582, found 227.0583.

(E)-1,2,4-Trifluoro-5-(2-nitrovinyl)benzene (9)

To a solution of MeOH, water and 2.5 N NaOH (30.0 g, 750.0 mmol, $V_{\text{MeOH}}/V_{\text{water}}/V_{\text{NaOH}} = 1/1/1$) was added to a solution of 2,4,5-trifluorobenzaldehyde (100.0 g, 624.6 mmol) and MeNO₂ (40.6 mL, 750.0 mmol) in MeOH (3 L) dropwise over 30-60 min, while the internal temperature was maintained between 5-10 °C. Then, the reaction solution was agitated another 30 min at 0-10 °C and added dropwise to a solution of ZnCl₂ (425.6 g, 3.12 mol) in conc. HCl (130 mL) and water (170 mL) for 2-4 h. After addition, the slurry was allowed to warm to ambient temperature and stirred for 1 h before filtration. The filter cake was washed with 40% MeOH aqueous solution and vacuum-dried at ambient temperature to give light yellow products **9** 99.0 g, 78% yield. ¹H NMR (500 MHz, DMSO-*d*₆): δ 7.95 (d, *J* = 14.0 Hz, 1H), 7.65 (d, *J* = 14.0 Hz, 1H), 7.39-7.33 (m, 1H), 7.12-7.05 (m, 1H). ¹³C NMR

(100 MHz, DMSO-*d*₆) δ 157.4, 152.3, 147.2, 139.7, 130.3, 118.2, 114.9, 107.0. LC-MS (ESI): [M+H]⁺ *m/z* 204.02, found 204.1.

6-Cyano-2,7-dimethoxy-1-naphthaldehyde (**11**)

To a solution of TiCl₄ (267.4 g, 1.41 mol) and 1,1-dichloromethyl methyl ether (80.5 g, 0.70 mmol) in dichloromethane (1 L) at 0 °C was added dropwise compound **12** (100 g, 0.47 mol) in DCM (2 L), maintaining the temperature below 5 °C. After completion of the addition, the reaction mixture was stirred at room temperature for 12-36 h. 1 N aqueous hydrochloric acid (835 mL) was added to quench the reaction. The organic layer was separated, and the aqueous layer was extracted with DCM. The organic layers were combined, washed with brine, dried over anhydrous Na₂SO₄. The solution was concentrated and purified by recrystallization to give the desired **11** 101.8 g, 90% yield. ¹H NMR (400 MHz, DMSO-*d*₆): δ 10.71 (s, 1H), 8.77 (s, 1H), 8.50 (s, 1H), 8.29 (d, *J* = 8.4 Hz, 1H), 7.55 (d, *J* = 8.8 Hz, 1H), 4.09 (s, 3H), 3.98 (s, 3H). ¹³C NMR (100 MHz, DMSO-*d*₆): δ 191.54, 166.43, 159.85, 137.79, 135.80, 135.39, 122.53, 116.22, 115.25, 111.17, 104.06, 101.80, 56.59, 56.26. LC-MS (ESI): [M+H]⁺ *m/z* 242.07, found 242.1; HRMS (EI): [M]⁺ *m/z* 241.0739, found 241.0742.

Ancillary Information

Supporting Information Availability

The Supporting Information is available free of charge on the ACS Publications website. Additional data and figures including FMO computational results, crystal X-ray structure data of **5**, RMSDs and single point energies (SPE) for the unbound, bound and minimum energy conformations of **5**, detailed dosing regimen of the long-term and body weight of the *db/db* mice, ¹H, ¹³C NMR and ESI-mass spectra of products, and HPLC of compound **5**, the molecular formula strings (CSV).

Abbreviations Used

DPP-4, dipeptidyl peptidase 4; T2DM, type 2 diabetes mellitus; HbA1c, hemoglobin A1c; GLP-1, glucagon-like peptide-1; SGLT2, sodium-glucose co-transporter 2; FMO, fragment molecular orbital; QM, quantum mechanics; MM, molecular mechanics; PIE, pair interaction energy; PIEDA, pair interaction energy decomposition analysis; ESP, electrostatic surface potential; RMSD, root mean square deviation; SPE, single point energy; OGTT, oral glucose tolerance test; FBG, fasting blood glucose; FBW, fasting body weight; PBW, postprandial body weight; hERG, human Ether-à-go-go-Related Gene

Accession Code

The coordinate and structure factor files for DPP-4 in complex with compound **5** is 5ZID. Authors will release the atomic coordinates and experimental data upon article publication.

Corresponding Author Information

*For Z.Z.: phone: +86-21-64250213; E-mail: zhjzhao@ecust.edu.cn

*For J.Y.L.: phone: +86-21-50801313; E-mail: jyli@simm.ac.cn

*For H.L.: Phone: +86-21-64250213; Email: hlli@ecust.edu.cn

Acknowledgments

This work was supported by the National Key Research and Development Program [2016YFA0502304 to H.L.]; the National Natural Science Foundation of China (grant 81825020 to H.L., 81803437 to S.L.); the National Science & Technology Major Project “Key New Drug Creation and Manufacturing Program”, China (Number: 2018ZX09711002); the Chinese Academy of Sciences (XDA12040204), the Fundamental Research Funds for the Central Universities, Special Program for Applied Research on Super Computation of the NSFC-Guangdong Joint Fund (the second phase) under Grant No.U1501501. Shiliang Li is sponsored by Shanghai Sailing Program (No. 18YF1405100). Honglin Li is also sponsored by National Program for Special Supports of Eminent Professionals and National Program for Support of Top-notch Young Professionals.

Conflict of interest

The authors declare no conflict of interest

Author contributions

S.L., C.Q. and S.C. performed research and drafted the manuscript, they contributed equally to the study. S.L. and H.L. designed the compound, performed the computational work and interpreted the data. C.Q., H.X., X.F., M.Z., and Z.Z. were in charge of compound synthesis. C.Q. obtained the single crystal X-ray structure of **5**. F.W. and L.Z. expressed and purified the DPP-4 protein, co-crystalized DPP-4 with inhibitor, and solved the X-ray crystal structure. J.W. and L.Z. performed surface plasmon resonance affinity measurements. M.S., D.L and C.X. tested *in vitro* and *in vivo* plasma DPP-4 inhibition. S.C., J.Y.L., Q.J. and R.W. performed the PK studies and investigated the acute and chronic antidiabetic efficacy of the compounds in mouse model systems. J.L and H.J. helped analyze the data and commented on the manuscript. H.L., Z.Z. and J.Y.L. conceived and supervised the overall project. All authors revised and approved the manuscript being submitted.

Reference

- (1) <http://www.idf.org/diabetesatlas> (accessed Jan. 31th, 2018).
- (2) Tahrani, A. A.; Bailey, C. J.; Del Prato, S.; Barnett, A. H. Management of type 2 diabetes: new and future developments in treatment. *Lancet* **2011**, *378* (9786), 182-197.
- (3) Reusch, J. E.; Manson, J. E. Management of type 2 diabetes in 2017: getting to goal. *JAMA* **2017**, *317* (10), 1015-1016.
- (4) Deacon, C. F.; Lebovitz, H. E. Comparative review of dipeptidyl peptidase-4 inhibitors and sulphonylureas. *Diabetes Obes. Metab.* **2016**, *18* (4), 333-347.
- (5) Waldrop, G.; Zhong, J.; Peters, M.; Rajagopalan, S. Incretin-based therapy for diabetes: what a cardiologist needs to know. *J. Am. Coll. Cardiol.* **2016**, *67* (12), 1488-1496.
- (6) Mulvihill, E. E.; Drucker, D. J. Pharmacology, physiology, and mechanisms of action of dipeptidyl peptidase-4 inhibitors. *Endocr. Rev.* **2014**, *35* (6), 992-1019.
- (7) Kushwaha, R. N.; Haq, W.; Katti, S. Discovery of 17 gliptins in 17-years of research for the treatment of type 2 diabetes: a synthetic overview. *Chem. Biol. Interface* **2014**, *4* (3), 137-162.
- (8) Scheen, A. J. Once-weekly DPP-4 inhibitors: do they meet an unmet need? *Lancet Diabetes Endo.* **2015**, *3* (3), 162-164.
- (9) Burness, C. B. Omarigliptin: first global approval. *Drugs* **2015**, *75* (16), 1947-1952.
- (10) McKeage, K. Trelagliptin: first global approval. *Drugs* **2015**, *75* (10), 1161-1164.
- (11) Yamada, T.; Shojima, N.; Noma, H.; Yamauchi, T.; Kadowaki, T. Weekly versus daily dipeptidyl peptidase 4 inhibitor therapy for type 2 diabetes: systematic review and meta-analysis. *Diabetes Care* **2018**, *41* (4), e52-e55.

1
2
3
4 (12) Suzuki, K.; Hasegawa, K.; Watanabe, M. Efficacy and patient satisfaction of
5
6 dipeptidyl peptidase-4 inhibitor after switching from once-daily DPP-4 inhibitor to
7
8 once-weekly regimen. *J. Clin. Med. Res.* **2018**, *10* (8), 641.

9
10
11 (13) Pipeline drug evidence review: Marizev (omarigliptin) vs. Januvia (sitagliptin).
12
13 Advera Health Analytics. **2016**, Available from: www.adverahealth.com.

14
15
16 (14) Stoimenis, D.; Karagiannis, T.; Katsoula, A.; Athanasiadou, E.; Kazakos, K.;
17
18 Bekiari, E.; Matthews, D. R.; Tsapas, A. Once-weekly dipeptidyl peptidase-4 inhibitors for
19
20 type 2 diabetes: a systematic review and meta-analysis. *Expert Opin. Pharmacother.* **2017**, *18*
21
22 (9), 843-851.

23
24
25 (15) Gantz, I.; Chen, M.; Suryawanshi, S.; Ntabadde, C.; Shah, S.; O'Neill, E. A.; Engel,
26
27 S. S.; Kaufman, K. D.; Lai, E. A randomized, placebo-controlled study of the cardiovascular
28
29 safety of the once-weekly DPP-4 inhibitor omarigliptin in patients with type 2 diabetes
30
31 mellitus. *Cardiovasc. Diabetol.* **2017**, *16* (1), 112.

32
33
34 (16) Khalse, M.; Bhargava, A. A review on cardiovascular outcome studies of dipeptidyl
35
36 peptidase-4 inhibitors. *Indian J. Endocrinol. Metab.* **2018**, *22* (5), 689.

37
38
39 (17) Inagaki, N.; Onouchi, H.; Maezawa, H.; Kuroda, S.; Kaku, K. Once-weekly
40
41 trelagliptin versus daily alogliptin in Japanese patients with type 2 diabetes: a randomised,
42
43 double-blind, phase 3, non-inferiority study. *Lancet Diabetes Endo.* **2015**, *3* (3), 191-197.

44
45
46 (18) Kaku, K. Safety evaluation of trelagliptin in the treatment of Japanese type 2
47
48 diabetes mellitus patients. *Expert Opin. Drug Saf.* **2017**, *16* (11), 1313-1322.

49
50
51 (19) FDA Drug Safety Communication: FDA adds warnings about heart failure risk to
52
53 labels of type 2 diabetes medicines containing saxagliptin and alogliptin. www.fda.gov,
54
55
56
57
58
59
60

1
2
3
4 accessed on 12 May 2017.
5

6 (20) Green, J. B.; Bethel, M. A.; Armstrong, P. W.; Buse, J. B.; Engel, S. S.; Garg, J.;
7 Josse, R.; Kaufman, K. D.; Koglin, J.; Korn, S. Effect of sitagliptin on cardiovascular
8 outcomes in type 2 diabetes. *N. Engl. J. Med.* **2015**, *373* (3), 232-242.
9
10
11

12 (21) Duan, L.; Rao, X.; Xia, C.; Rajagopalan, S.; Zhong, J. The regulatory role of DPP4
13 in atherosclerotic disease. *Cardiovasc. Diabetol.* **2017**, *16* (1), 76.
14
15
16

17 (22) Li, S.; Xu, H.; Cui, S.; Wu, F.; Zhang, Y.; Su, M.; Gong, Y.; Qiu, S.; Jiao, Q.; Qin,
18 C.; Shan, J.; Zhang, M.; Wang, J.; Yin, Q.; Xu, M.; Liu, X.; Wang, R.; Zhu, L.; Li, J.; Xu, Y.;
19 Jiang, H.; Zhao, Z.; Li, J.; Li, H. Discovery and rational design of natural-product-derived
20 2-phenyl-3, 4-dihydro-2 H-benzo [f] chromen-3-amine analogs as novel and potent dipeptidyl
21 peptidase 4 (DPP-4) inhibitors for the treatment of type 2 diabetes. *J. Med. Chem.* **2016**, *59*
22 (14), 6772-6790.
23
24
25
26
27
28
29
30
31
32
33

34 (23) Walkup, G. K.; You, Z.; Ross, P. L.; Allen, E. K.; Daryaei, F.; Hale, M. R.;
35 O'Donnell, J.; Ehmann, D. E.; Schuck, V. J.; Buurman, E. T. Translating slow-binding
36 inhibition kinetics into cellular and in vivo effects. *Nat. Chem. Biol.* **2015**, *11* (6), 416-423.
37
38
39
40
41
42

43 (24) Vauquelin, G. Effects of target binding kinetics on in vivo drug efficacy: koff, kon
44 and rebinding. *Br. J. Pharmacol.* **2016**, *173* (15), 2319-2334.
45
46
47

48 (25) Copeland, R. A. The drug-target residence time model: a 10-year retrospective. *Nat.*
49 *Rev. Drug Discov.* **2016**, *15*, 87-95.
50
51
52

53 (26) Dahl, G.; Akerud, T. Pharmacokinetics and the drug-target residence time concept.
54 *Drug Discov. Today* **2013**, *18* (15-16), 697-707.
55
56
57

58 (27) Tonge, P. J. Drug-target kinetics in drug discovery. *ACS Chem. Neurosci.* **2017**, *9*
59
60

1
2
3
4 (1), 29-39.
5

6 (28) Fedorov, D. G.; Kitaura, K. Pair interaction energy decomposition analysis. *J.*
7
8
9 *Comput. Chem.* **2007**, *28* (1), 222-237.
10

11 (29) Fedorov, D. G.; Nagata, T.; Kitaura, K. Exploring chemistry with the fragment
12
13
14
15
16 molecular orbital method. *PCCP* **2012**, *14* (21), 7562-7577.

17 (30) Fedorov, D. G.; Kitaura, K. Extending the power of quantum chemistry to large
18
19
20
21
22
23
24 systems with the fragment molecular orbital method. *J. Phys. Chem. A* **2007**, *111* (30),
25
26
27
28
29
30
31 6904-6914.

32 (31) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment molecular
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60 orbital method: an approximate computational method for large molecules. *Chem. Phys. Lett.*
1999, *313* (3-4), 701-706.

(32) Vauquelin, G.; Charlton, S. J. Long-lasting target binding and rebinding as
mechanisms to prolong in vivo drug action. *Br. J. Pharmacol.* **2010**, *161* (3), 488-508.

(33) de Witte, W.; Vauquelin, G.; van der Graaf, P.; de Lange, E. The influence of drug
distribution and drug-target binding on target occupancy: The rate-limiting step
approximation. *Eur. J. Pharm. Sci.* **2017**, *109*, S83-S89.

(34) de Witte, W. E.; Danhof, M.; van der Graaf, P. H.; de Lange, E. C. In vivo target
residence time and kinetic selectivity: the association rate constant as determinant. *Trends*
Pharmacol. Sci. **2016**, *37* (10), 831-842.

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

(36) Biftu, T.; Sinha-Roy, R.; Chen, P.; Qian, X. X.; Feng, D.; Kuethe, J. T.; Scapin, G.;

1
2
3
4 Gao, Y. D.; Tan, Y. W.; Krueger, D.; Bak, A.; Eiermann, G.; He, J. F.; Cox, J.; Hicks, J.;
5
6 Lyons, K.; He, H. B.; Salituro, G.; Tong, S. R.; Patel, S.; Doss, G.; Petrov, A.; Wu, J.; Xu, S.
7
8 S.; Sewall, C.; Zhang, X. P.; Zhang, B.; Thornberry, N. A.; Weber, A. E. Omarigliptin
9
10 (MK-3102): a novel long-acting DPP-4 inhibitor for once-weekly treatment of type 2 diabetes.
11
12 *J. Med. Chem.* **2014**, *57* (8), 3205-3212.
13
14

15
16
17 (37) Kim, D.; Wang, L.; Beconi, M.; Eiermann, G. J.; Fisher, M. H.; He, H.; Hickey, G.
18
19 J.; Kowalchick, J. E.; Leiting, B.; Lyons, K.; Marsilio, F.; McCann, M. E.; Patel, R. A.;
20
21 Petrov, A.; Scapin, G.; Patel, S. B.; Roy, R. S.; Wu, J. K.; Wyvratt, M. J.; Zhang, B. B.; Zhu,
22
23 L.; Thornberry, N. A.; Weber, A. E.
24
25 (2R)-4-oxo-4-[3-(trifluoromethyl)-5,6-dihydro[1,2,4]triazolo[4,3-a]pyrazin-7(8H)-
26
27 yl]-1-(2,4,5-trifluorophenyl)butan-2-amine: a potent, orally active dipeptidyl peptidase IV
28
29 inhibitor for the treatment of type 2 diabetes. *J. Med. Chem.* **2005**, *48* (1), 141-151.
30
31
32
33

34
35 (38) Pratley, R. E.; Salsali, A. Inhibition of DPP-4: a new therapeutic approach for the
36
37 treatment of type 2 diabetes. *Curr. Med. Res. Opin.* **2007**, *23* (4), 919-931.
38
39

40
41 (39) Liu, X.; Zhang, L. N.; Feng, Y.; Zhang, L.; Qu, H.; Cao, G. Q.; Leng, Y. Acute and
42
43 chronic administration of SHR117887, a novel and specific dipeptidyl peptidase-4 inhibitor,
44
45 improves metabolic control in diabetic rodent models. *Acta Pharmacol Sin* **2012**, *33* (8),
46
47 1013-22.
48
49

50
51 (40) Nojima, H.; Kanou, K.; Terashi, G.; Takeda-Shitaka, M.; Inoue, G.; Atsuda, K.; Itoh,
52
53 C.; Iguchi, C.; Matsubara, H. Comprehensive analysis of the co-structures of dipeptidyl
54
55 peptidase IV and its inhibitor. *BMC Struct. Biol.* **2016**, *16* (1), 11.
56
57

58
59 (41) Nabeno, M.; Akahoshi, F.; Kishida, H.; Miyaguchi, I.; Tanaka, Y.; Ishii, S.;
60

1
2
3
4 Kadowaki, T. A comparative study of the binding modes of recently launched dipeptidyl
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Kadowaki, T. A comparative study of the binding modes of recently launched dipeptidyl
peptidase IV inhibitors in the active site. *Biochem. Biophys. Res. Commun.* **2013**, *434* (2),
191-196.

(42) Mazanetz, M. P.; Ichihara, O.; Law, R. J.; Whittaker, M. Prediction of
cyclin-dependent kinase 2 inhibitor potency using the fragment molecular orbital method. *J.*
Cheminform. **2011**, *3* (1), 2.

(43) Heifetz, A.; Trani, G.; Aldeghi, M.; MacKinnon, C. H.; McEwan, P. A.; Brookfield,
F. A.; Chudyk, E. I.; Bodkin, M.; Pei, Z.; Burch, J. D. Fragment molecular orbital method
applied to lead optimization of novel interleukin-2 inducible T-cell kinase (ITK) inhibitors. *J.*
Med. Chem. **2016**, *59* (9), 4352-4363.

(44) Heifetz, A.; Chudyk, E. I.; Gleave, L.; Aldeghi, M.; Cherezov, V.; Fedorov, D. G.;
Biggin, P. C.; Bodkin, M. J. The Fragment molecular orbital method reveals new insight into
the chemical nature of GPCR-ligand interactions. *J. Chem. Inf. Model.* **2015**, *56* (1), 159-172.

(45) Sriwilaijaroen, N.; Magesh, S.; Imamura, A.; Ando, H.; Ishida, H.; Sakai, M.;
Ishitsubo, E.; Hori, T.; Moriya, S.; Ishikawa, T. A novel potent and highly specific inhibitor
against influenza viral N1-N9 neuraminidases: insight into neuraminidase-inhibitor
interactions. *J. Med. Chem.* **2016**, *59* (10), 4563-4577.

(46) Liu, X.; Jiang, H.; Li, H. SHAFTS: a hybrid approach for 3D molecular similarity
calculation. 1. Method and assessment of virtual screening. *J. Chem. Inf. Model.* **2011**, *51* (9),
2372-2385.

(47) Lu, W.; Liu, X.; Cao, X.; Xue, M.; Liu, K.; Zhao, Z.; Shen, X.; Jiang, H.; Xu, Y.;
Huang, J.; Li, H. SHAFTS: a hybrid approach for 3D molecular similarity calculation. 2.

1
2
3
4 Prospective case study in the discovery of diverse p90 ribosomal S6 protein kinase 2
5
6 inhibitors to suppress cell migration. *J. Med. Chem.* **2011**, *54* (10), 3564-3574.
7

8
9 (48) Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen,
10
11 J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. General atomic and molecular
12
13 electronic structure system. *J. Comput. Chem.* **1993**, *14* (11), 1347-1363.
14

15
16 (49) Fedorov, D. G. The fragment molecular orbital method: theoretical development,
17
18 implementation in GAMESS, and applications. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*
19
20 **2017**, *7* (6), e1322.
21

22
23 (50) Fedorov, D. G.; Kitaura, K. Second order Møller-Plesset perturbation theory based
24
25 upon the fragment molecular orbital method. *J. Chem. Phys.* **2004**, *121* (6), 2483-2490.
26
27

28
29 (51) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation
30
31 methods: a route to accurate calculations on large systems. *Chem. Rev.* **2011**, *112* (1),
32
33 632-672.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table of Contents graphic

