

RESEARCH ARTICLE

Open Access



# Comprehensive analysis of the Co-structures of dipeptidyl peptidase IV and its inhibitor

Hiroyuki Nojima<sup>1\*</sup>, Kazuhiko Kanou<sup>1,2</sup>, Genki Terashi<sup>1</sup>, Mayuko Takeda-Shitaka<sup>1</sup>, Gaku Inoue<sup>1</sup>, Koichiro Atsuda<sup>1</sup>, Chihiro Itoh<sup>1</sup>, Chie Iguchi<sup>1</sup> and Hajime Matsubara<sup>1</sup>

## Abstract

**Background:** We comprehensively analyzed X-ray cocrystal structures of dipeptidyl peptidase IV (DPP-4) and its inhibitor to clarify whether DPP-4 alters its general or partial structure according to the inhibitor used and whether DPP-4 has a common rule for inhibitor binding.

**Results:** All the main and side chains in the inhibitor binding area were minimally altered, except for a few side chains, despite binding to inhibitors of various shapes. Some residues (Arg125, Glu205, Glu206, Tyr662 and Asn710) in the area had binding modes to fix a specific atom of inhibitor to a particular spatial position in DPP-4. We found two specific water molecules that were common to 92 DPP-4 structures. The two water molecules were close to many inhibitors, and seemed to play two roles: maintaining the orientation of the Glu205 and Glu206 side chains through a network via the water molecules, and arranging the inhibitor appropriately at the S2 subsite.

**Conclusions:** Our study based on high-quality resources may provide a necessary minimum consensus to help in the discovery of a novel DPP-4 inhibitor that is commercially useful.

**Keywords:** Dipeptidyl peptidase IV, DPP-4 inhibitor, Inhibitory activity, Cocrystal structure, Water molecule, In silico screening

## Background

Incretin is an endogenous gut hormone that is useful in treating patients with type 2 diabetes [1]. Incretin is secreted from the digestive tract with dietary intake [2] and acts on pancreatic  $\beta$ -cells to stimulate insulin secretion [1, 3]. Sulfonylurea, a traditional hypoglycemic drug, promotes insulin secretion regardless of the blood glucose level; because of this, it risks eliciting serious hypoglycemia. In contrast, an antidiabetic drug that uses incretin as a mediator is expected to reduce the risk of hypoglycemia because stimulation of insulin secretion by incretin depends on the blood glucose level [4].

Glucagon-like peptide-1 (GLP-1) is an incretin with strong insulin secretion effect [5]. The active form of GLP-1 comprises 30 amino acids [GLP-1-(7–36)NH<sub>2</sub> or

GLP-1-(7–37)] [6], but it has a short half-life of only 2 min because two residues (His-Ala) on the N-terminus of the active form are removed by dipeptidyl peptidase IV (DPP-4) [7]. Currently, two different types of drugs are in clinical use to target GLP-1. The first type is a GLP-1 analog, which has a longer half-life than active endogenous GLP-1; examples of this type include liraglutide (half-life = 13 h) [8], exenatide (half-life = 1.3–1.6 h) [9] and lixisenatide (half-life = approximately 2 h) [10]. The second type is a DPP-4 inhibitor, which prolongs the half-life of active endogenous GLP-1 by inhibiting DPP-4.

DPP-4 has strong protease activity against polypeptides that have an alanine or proline as a second N-terminal residue [11]. DPP-4 inhibitor development began with dipeptide structures that included alanine or proline as the base. Currently, nine DPP-4 inhibitors are marketed in many countries: sitagliptin [2], vildagliptin [12, 13], alogliptin [14, 15], linagliptin [16], anagliptin [17, 18], teneligliptin [19], saxagliptin

\* Correspondence: nojimah@pharm.kitasato-u.ac.jp

<sup>1</sup>School of Pharmacy, Kitasato University, 5-9-1 ShirokaneMinato-ku, Tokyo 108-8641, Japan

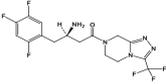
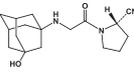
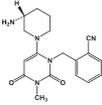
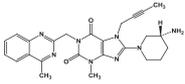
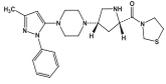
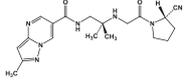
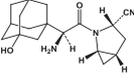
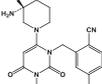
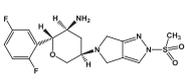
Full list of author information is available at the end of the article

[20], trelagliptin [15, 21] and omarigliptin [22]. However, these nine drugs have different structures, with the exception of vildagliptin and saxagliptin, or alogliptin and trelagliptin [23–25] (Table 1).

An explanation for why DPP-4 can accept inhibitors of various shapes is that DPP-4 has a large cavity

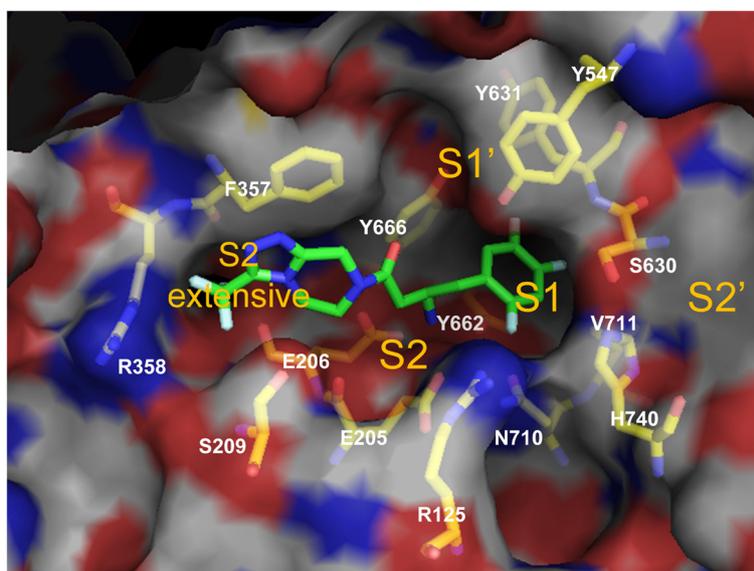
(diameter  $\geq 20$  Å) [26], and these inhibitors may be allowed to approach the active center of DPP-4. In addition, DPP-4 has multiple binding subsites known as the S1, S2, S1', S2' and S2 extensive subsites (Fig. 1) [13]. Of the commercial drugs, vildagliptin and saxagliptin bind to the S1 and S2 subsites, alogliptin, linagliptin and

**Table 1** Structural similarity of commercial DPP-4 inhibitors (Tanimoto coefficient<sup>a</sup>)

	Sitagliptin	Vildagliptin	Alogliptin	Linagliptin	Teneligliptin	Anagliptin	Saxagliptin	Trelagliptin	Omarigliptin
 Sitagliptin (IC <sub>50</sub> <sup>b</sup> : 18 nM)	-								
 Vildagliptin (IC <sub>50</sub> : 3.5 nM)	0.351	-							
 Alogliptin (IC <sub>50</sub> : 7 nM)	0.325	0.424	-						
 Linagliptin (IC <sub>50</sub> : 1 nM)	0.286	0.239	0.364	-					
 Teneligliptin (IC <sub>50</sub> : 0.37 nM)	0.261	0.300	0.341	0.413	-				
 Anagliptin (IC <sub>50</sub> : 3.8 nM)	0.244	0.429	0.233	0.286	0.349	-			
 Saxagliptin (Ki: 0.6 nM)	0.378	<b>0.800</b>	0.371	0.289	0.233	0.308	-		
 Trelagliptin (IC <sub>50</sub> : 4 nM)	0.350	0.412	<b>0.962</b>	0.356	0.366	0.227	0.361	-	
 Omarigliptin (IC <sub>50</sub> : 1.6 nM, Ki: 0.8 nM)	0.410	0.324	0.368	0.292	0.326	0.146	0.250	0.395	-

<sup>a</sup> Tanimoto coefficients were calculated by chemical structure comparison using the build-up algorithm [23]. They range from 0 to 1. When 0.8 or higher, two structures are evaluated as similar (bold)

<sup>b</sup> IC<sub>50</sub> and Ki were quoted from the Web Server "The Binding Database", <http://www.bindingdb.org/bind/> [24] or the Web Server "PDB bind", <http://www.pdbbind.org.cn/> [25]



**Fig. 1** Inhibitor binding area of DPP-4. Representative image from the cocrystal structure of sitagliptin and DPP-4 (PDB ID: 1X70). The carbon skeleton of sitagliptin is represented by green stick. The carbon skeleton of 14 residues is labeled and represented by yellow stick. Val656 and Trp659 are positioned on the opposite side of view in this Figure; thus, they are not shown. O, N, and halogen atoms are labeled in red, blue and light blue sticks, respectively. The subsites, which are directly involved with binding to inhibitors (S1, S1', S2, S2' and S2 extensive), are labeled in orange

possibly trelagliptin bind to the S1' and/or S2' subsites in addition to the S1 and S2 subsites, while sitagliptin, anagliptin, teneligliptin and omarigliptin bind to the S1, S2 and S2 extensive subsites. The commercial drugs efficiently match the energy in these subsites and, in this manner, probably attain high DPP-4 inhibitory activity.

It is not known if there are other causes for DPP-4 binding to inhibitors of various shapes. For example, does DPP-4 alter its general or partial structure according to the inhibitor used? Or, does DPP-4 have a common rule for inhibitor binding? To answer these questions, we comprehensively analyzed X-ray cocrystal structures of DPP-4 and its inhibitor. All the main and side chains in the inhibitor binding area were minimally altered, except for some side chains, despite binding to inhibitors of various shapes. Some residues in the area had binding modes to fix a specific atom of inhibitor to a particular spatial position in DPP-4. We found two specific water molecules that were common to many DPP-4 structures. The two water molecules were close to many inhibitors, and seemed to be related to inhibitor binding. This information may provide a necessary minimum consensus to help in the discovery of a novel DPP-4 inhibitor that is commercially useful.

## Methods

### Data collection

We collected X-ray cocrystal structures of human DPP-4 and its inhibitor that were registered with the

Protein Data Bank (PDB) [27] until 2015. Sixty-eight PDB codes that had a resolution of less than 3 Å were used (Additional file 1: Figure S1). Most of the PDBs had a crystallization temperature that ranged from 277 K to 298 K and an X-ray diffraction-measured temperature range from 90 K to 120 K. However, the X-ray diffraction of only five PDBs (PDB ID: 2AJL, 2I03, 2I78, 2OLE and 3EIO) was measured at a high temperature (in the range of 200 K to 298 K) (Additional file 2: Table S1).

One, two or four DPP-4 molecules are included per one PDB code. The DPP-4 molecule that had an inhibitor bound to the DPP-4 active center and that had less than six disordered residues was selected from each PDB code. We defined the coordinates of one DPP-4 molecule (724 residues: residue 41–764) with one inhibitor on the active center and water O atoms within 4 Å from the DPP-4 molecule as one unit (we will discuss the distance between heavy atoms). Ultimately, there were 147 inhibitor-bound units identified from the 68 PDBs (i.e., 68 kinds of inhibitors). To compare and evaluate the inhibitor-bound units, we also collected X-ray crystal structures of inhibitor-free human DPP-4 that had a resolution of less than 3 Å. Eight inhibitor-free units were identified from four PDB codes (PDB ID: 1J2E, 1NU6, 1PFQ and 1TK3) (Additional file 2: Table S1). These units were used for the procedure discussed below.

### Defining the DPP-4 inhibitor binding area

Generally, some kinds of interactions (e.g., hydrogen bond, electrostatic interaction, hydrophobic interaction and  $\pi$ - $\pi$  stacking effect) are considered to occur between two heavy atoms that are close to each other (less than c.a. 4–5 Å). In DPP-4, 13 residues (Arg125, Glu205, Glu206, Phe357, Tyr547, Ser630, Tyr631, Val656, Tyr662, Tyr666, Asn710, Val711 and His740) were close to (<4 Å) at least 48 out of the 68 inhibitors (>70 %) and three residues (Ser209, Arg358 and Trp659) were close to (<4 Å) at least 21 out of the 68 inhibitors (>30 %). We defined the above 16 residues as the DPP-4 inhibitor binding area (Fig. 1).

### C $\alpha$ atom variation between units

We calculated C $\alpha$  atom variation between units as follows:

[Step 1] From the above-mentioned units, two units were selected.

[Step 2] The two units were superimposed so that the root mean square deviation (RMSD) targeting C $\alpha$  atoms of the DPP-4 molecule (residue 41–764) would be minimized. The RMSD is generally defined by the following Eq. (1):

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_i^N x_i^2} \quad (1)$$

where  $x_i$  represents the distance of the  $i^{\text{th}}$  atom between the two units, and  $N$  represents the number of equivalent atom pairs. In this procedure, only the C $\alpha$  atom was applied to Eq. (1). The minimum RMSD indicates the global deviation of DPP-4 structure between the two units (the minimum RMSD value will be referred to as RMSD). A graphics software program, PyMOL (Schrödinger, Inc., New York, NY, USA), was used for superposition [28].

[Step 3] Steps 1 and 2 were conducted for every combination. When targeting 147 inhibitor-bound units, there are 10,731 combinations.

[Step 4] After optimum superposition for all combinations was achieved, the average distance of the  $i^{\text{th}}$  C $\alpha$  atom between two units was calculated using Eq. 2, and defined as the  $i^{\text{th}}$  C $\alpha$  atom variation between units.

$$[\text{Variation between units}]_i = \frac{1}{M} \sum_k^M \delta_{ik} \quad (2)$$

where  $\delta_{ik}$  is the distance of the  $i^{\text{th}}$  atom for the  $k^{\text{th}}$  combination between two units, and  $M$  represents the number of combinations. When targeting 147 inhibitor-bound units,  $M$  is 10,731.

### Side chain variation between units

Side chain variation between units was calculated as follows:

[Step 1] From the above-mentioned units, two units were selected.

[Step 2] The two units were superimposed so that the RMSD of main chain (N, C $\alpha$  and C atoms) of the  $i^{\text{th}}$  residue would be minimized (Eq. 1 was used). The minimum RMSD calculated here represents the main chain deviation of the  $i^{\text{th}}$  residue between the two units.

[Step 3] Steps 1 and 2 were conducted for every combination.

[Step 4] After superposition for all combinations, the average distance of a specific side chain atom (e.g., the serine O $\gamma$  atom) between two units was calculated using Eq. 2 from the previous procedure. This was defined as the side chain variation of the  $i^{\text{th}}$  residue between units.

### Calculating the exposed surface area

The exposed surface area per one residue of one unit was calculated. In the calculation, water O atoms and the inhibitor were excluded from the unit. The exposed surface area was determined using the Web Server GETAREA (Sealy Center for Structural Biology, University of Texas Medical Branch, Galveston, TX, USA) [29]. For each residue, an average of 147 inhibitor-bound units was calculated (this value is called the exposed surface area).

## Results and discussion

### C $\alpha$ atom variation

The average RMSD targeting C $\alpha$  atoms between 147 inhibitor-bound units was 0.48 Å and the maximum RMSD was 0.98 Å. Considering that the minimum resolution of the adopted structures, that is the minimum coordinate error of the structures used, is 1.62 Å (PDB ID: 4A5S, [30]) and that DPP-4 is a large molecule composed of more than 700 residues, we suggest that even the maximum RMSD is below the measurement error range and the units examined are all similar. Keedy et al. reports that the crystal structure of a protein is affected more by cryocooling than by the lab performing the experiment [31]. Eight units (PDB ID: 2AJL\_I, 2AJL\_J, 2I03\_B, 2I78\_B, 2OLE\_A, 2OLE\_B, 3EIO\_A and 3EIO\_B) were measured in X-ray diffraction at a higher temperature (in the range of 200 K to 298 K) than the other inhibitor-bound units (Additional file 2: Table S1), but the high temperature-measured units were not large in the average RMSD compared with the other units (in the range from 0.46 Å to 0.56 Å). This result suggests that DPP-4 global structure is not changed in the temperature range from 90 to 298 K.

The RMSD measures global deviation between units, but cannot identify partial deviation within molecule. Thus, each C $\alpha$  atom variation between units was calculated and presented in the order of the exposed surface area (Fig. 2). Generally, in a large molecule, the surface (outside) is more variegated than the inside because the former is more affected by the external environment (e.g., a crystal forming condition or measurement temperature) than the latter. The larger the exposed surface area, the more frequently C $\alpha$  atoms with a large variation were found in the DPP-4 structure. However, all residues in the inhibitor binding area had a less variation than the mean variation of the 724 residues (0.37 Å, Fig. 2), indicating that the main chain structure of the inhibitor binding area was only slightly altered according to the inhibitor used.

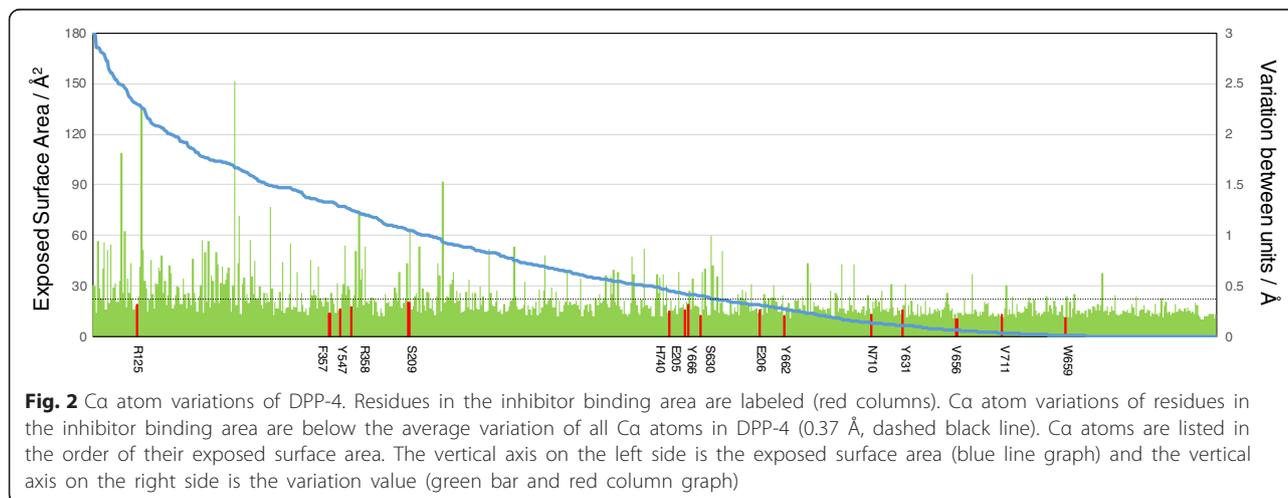
### Side chain variation

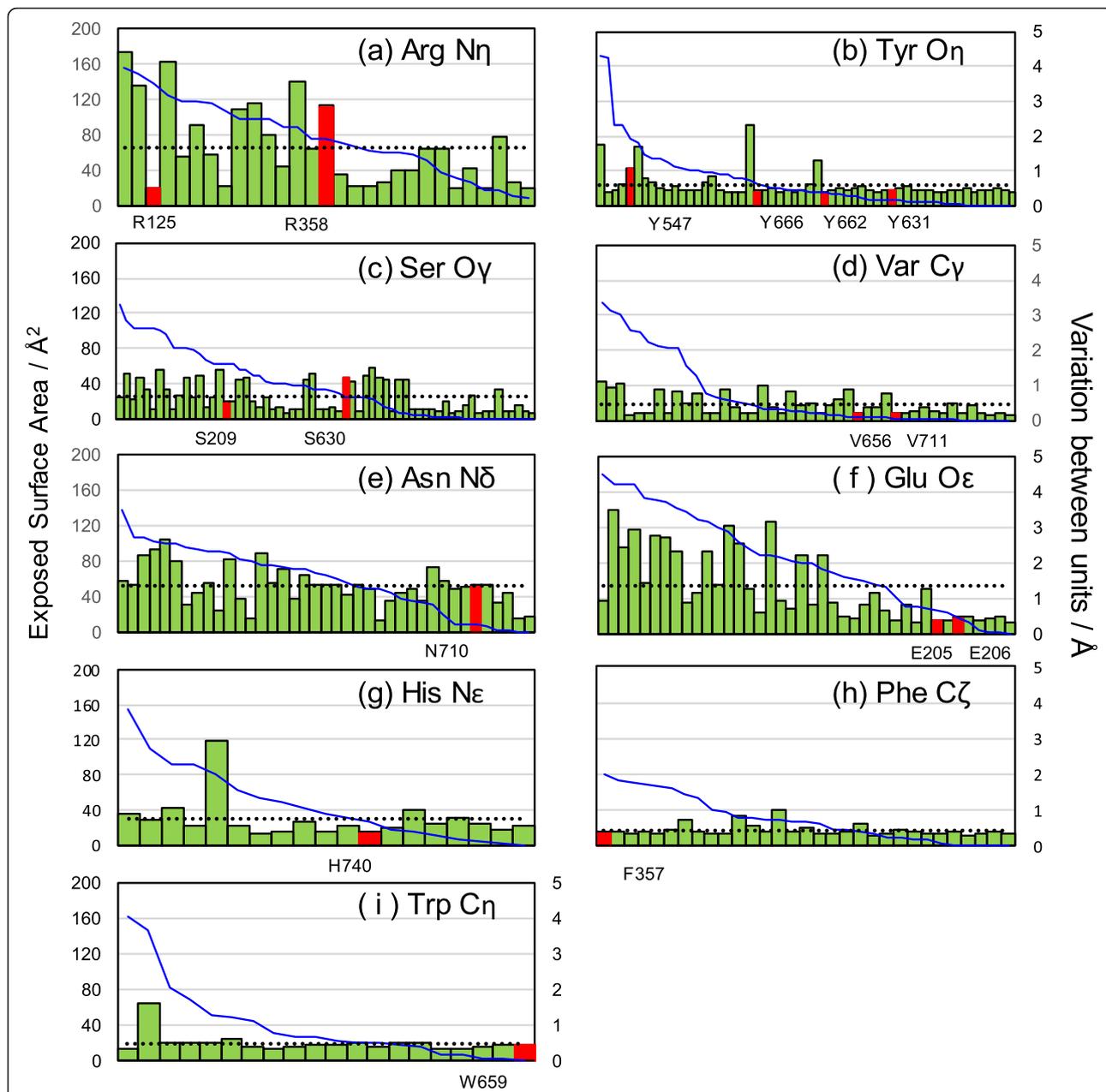
Side chain variation was classified for each amino acid class and presented in the order of the exposed surface area (Fig. 3). We also visually observed the superposition of 147 inhibitor-bound units (Fig. 4). In the inhibitor binding area, the side chains of Arg358, Tyr547 and Ser630 had a larger variation compared with the average of the equivalent amino acids (Fig. 3a, b and c).

The Arg358 side chains were oriented in a disorderly manner (Fig. 4a). This disorder was found in the inhibitor-free units (Additional file 3: Figure S2a). Arg358 constitutes a part of the S2 extensive subsite, and out of 68, 21 inhibitors (out of 147 inhibitor-bound units, 34 units) were close to this residue (<4 Å). However, there was no trend in the orientation of the Arg358 side chains, even when focusing on only the 34 units (the N $\eta$  atoms are colored cyan in Fig. 4a). These results indicate that the Arg358 side chain has no specific orientation regardless of the presence/absence of

inhibitor. To inquire into the cause of this large variation, we compared Arg358 with Arg125 whose side chain had the minimum variation (Fig. 4a and b). Arg125 had concentrated distributions of some water O atoms surrounding the whole of residue (Fig. 4b), whereas Arg358 had no concentrated distribution of water O atom surrounding the N $\eta$  atoms (Fig. 4a). The exposed surface area of Arg125 is larger than that of Arg358 at first sight, but the Arg125 side chain may be fixed because it is surrounded by some fixed hydrated water molecules. On the other hand, the perimeter of the Arg358 N $\eta$  atoms is free from hydrated water, and therefore the orientation of the Arg358 side chain may be flexible.

The Tyr547 benzene rings had their orientation depolarized (Fig. 4c and d). Sheehan et al. reports that the Tyr547  $\chi^1$  dihedral angle changes by 70° between the two orientations [32]. The Tyr547 benzene rings of the inhibitor-free units showed only one direction (Additional file 3: Figure S2b), and this direction was similar to that of one group of the inhibitor-bound units (Fig. 4c: the first group, the O $\eta$  atoms are colored red). The other group (Fig. 4d: the second group, the O $\eta$  atoms are colored cyan) had aromatic ring of the inhibitors stacked on the Tyr547 benzene ring. In the commercial drugs, sitagliptin, saxagliptin, trelagliptin, vildagliptin, anagliptin and omarigliptin were the first group, whereas linagliptin and alogliptin were the second group. These results suggest that the Tyr547 benzene ring could shift to a different direction from the original direction observed for the inhibitor-free units to obtain  $\pi$ - $\pi$  stacking interaction with inhibitor. The  $\pi$ - $\pi$  stacking interaction of the Tyr547 benzene ring has been reported by many studies [13–16, 30, 32–39]. However, the Tyr547 hydroxyl group in the original direction sometimes electrostatically interacts with the inhibitor's polar group or with hydrated water [2, 20, 33, 40, 41].

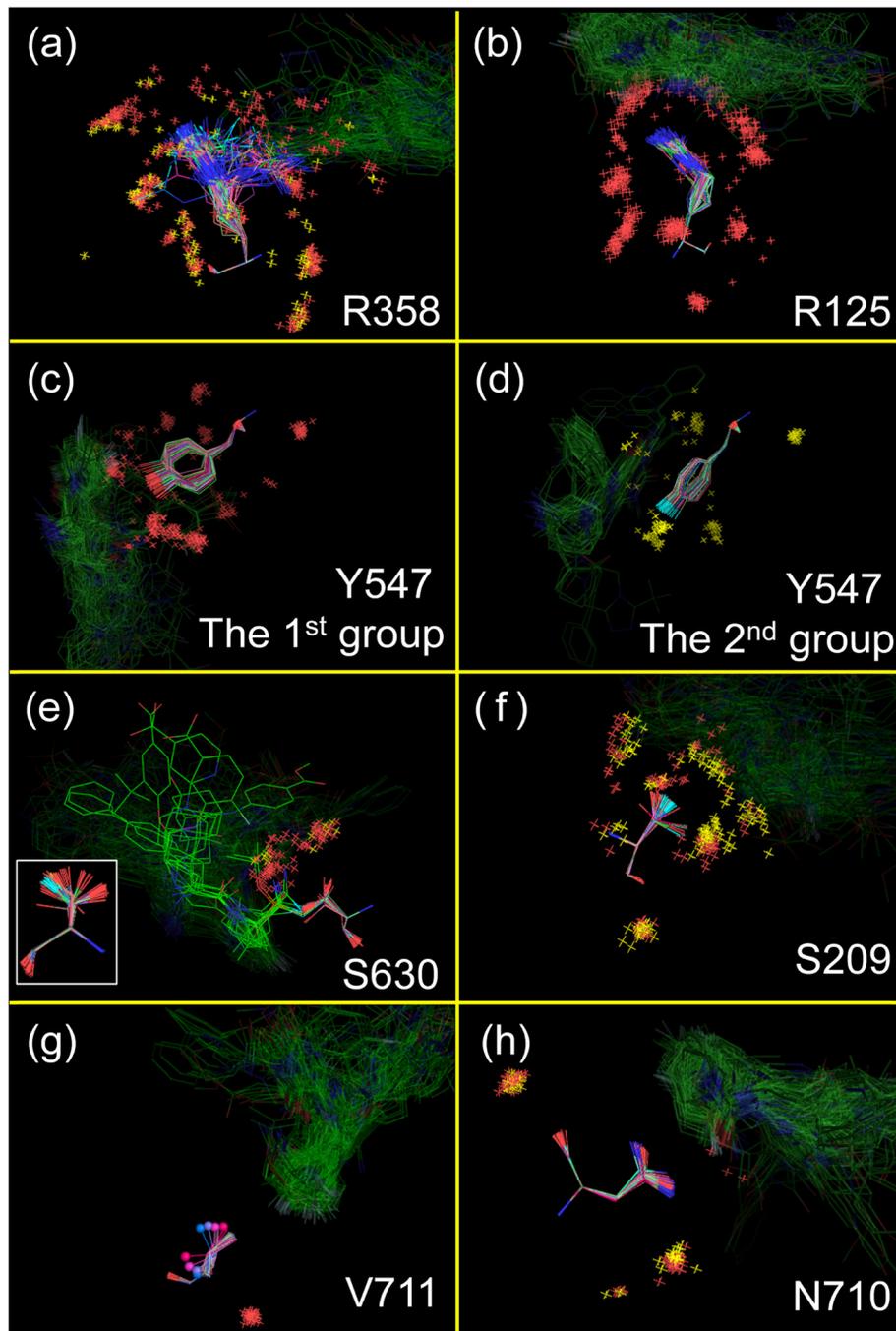




**Fig. 3** Side chain variations of DPP-4. Sixteen residues in the inhibitor binding area are labeled (red columns). Thirteen residues (Arg125, Tyr 631, Tyr 662, Tyr 666, Ser209, Val656, Val711, Asn710, Glu205, Glu206, His740 and Phe357 and Trp659) have their side chain variations below the average variation of the equivalent amino acids (dashed black line). Specific atoms that are positioned furthest from the main chain were calculated to be the variation. The graph shows the variations for each amino acid and they are listed in the order of the exposed surface area. The vertical axis on the left is the exposed surface area (blue line graph) and the vertical axis on the right is the variation value (green bar and red column graph). The variations of the two Arg N $\eta$  atoms, the two Glu O $\epsilon$  atoms and the two Val C $\gamma$  atoms are averaged, respectively. **a** Arg N $\eta$  atom, **b** Tyr O $\eta$  atom, **c** Ser O $\gamma$  atom, **d** Val O $\gamma$  atom, **e** Asn N $\delta$  atom, **f** Glu O $\epsilon$  atom, **g** His N $\epsilon$  atom, **h** Phe C $\zeta$  atom and **i** Trp C $\eta$  atom

The S630 O $\gamma$  atoms were oriented in all directions (Fig. 4e). This disorder was found in the inhibitor-free units (Additional file 3: Figure S2c). Ser630 is the active center of DPP-4 and was positioned within 4 Å from all 68 inhibitors. Some inhibitors have a cyanopyrrolidine, in which the nitrile C atom covalently bonds to the

Ser630 O $\gamma$  atom. In this case, the Ser630 O $\gamma$  atom was naturally oriented toward the inhibitor and very close to the inhibitor nitrile C atom (<3 Å, the O $\gamma$  atom is colored cyan in Fig. 4e) (PDB ID: 2AJL [42], 2G5P, 2G5T, 2G63 [33], 2I03 [43], 3BJM (saxagliptin) [20] and 3W2T (vildagliptin) [13]). However, for inhibitors that



**Fig. 4** Superposition of some residues in the inhibitor binding area. In each residue, 147 inhibitor-bound units are superimposed so that the RMSD targeting main chain (N, Ca and C atoms) would be minimized based on a specific inhibitor-bound unit. Excluding some exceptions, O and N atoms are colored red and blue, respectively, and the carbon skeleton of inhibitors is shadowed in green. **a** Arg358: The units close to the inhibitor ( $<4 \text{ \AA}$ ) have the  $N_{\eta}$  atom colored cyan and the water O atom colored yellow. **b** Arg125. **c** and **d** Tyr547: The Tyr547 side chains were divided into **c** the first and **d** the second group. In the second group, the  $O_{\eta}$  atom is colored cyan and the water O atom is colored yellow. **e** Ser630: The units in which inhibitor forms a covalent bond have the  $O_{\gamma}$  atom colored cyan, the inhibitor drawn by a line and the water O atom colored yellow. The image of the different angle is drawn at the lower left in this Figure. **f** Ser209: Some units close to the inhibitor ( $<4 \text{ \AA}$ ) have the  $O_{\gamma}$  atom colored cyan and the water O atom colored yellow. **g** Val711: The units in which the  $C_{\gamma}$  atoms have an opposite direction to the other units have the  $C_{\gamma}$  atom marked by a "sphere". **h** Asn710: Some units in which the  $O_{\delta}$  and  $N_{\delta}$  atoms are switched to each other have the water O atom colored yellow

do not require a covalent bond, the Ser630 O $\gamma$  atom seems to be oriented in a disorderly manner. This result suggests that the important thing for DPP-4 inhibitor may be in preventing GLP-1 from approaching Ser630 spatially rather than in deactivating Ser630 directly. Saxagliptin and vildagliptin are commercial covalent bond drugs, but they do not always have a higher potency than the other noncovalent bond commercial drugs (Table 1). This may be because DPP-4 itself slowly hydrolyzes their covalent bond and finally deactivates the covalent bond drugs [44].

The side chains of Ser209, Asn710 and Val711 did not have a significantly large variation compared with the average of the equivalent amino acids (Fig. 3c, d and e), but the superposition of the equivalent amino acids revealed irregular orientational alternatives (Fig. 4f, g and h). In the inhibitor-free units, the Ser209 and Val711 side chains showed the regular orientations (Additional file 3: Figure S2d and e). The Ser209 side chains had no trend when “close to inhibitor” (<4 Å, the O $\gamma$  atoms are colored cyan in Fig. 4f) and “not close to inhibitor” (the O $\gamma$  atoms are colored red in Fig. 4f) were compared. Val711 had only four units (3O95\_B, 3O95\_D, 3O9V\_B and 3O9V\_D) in which the C $\gamma$  atoms were irregularly oriented against the other inhibitor-bound units (the C $\gamma$  atoms are marked by sphere in Fig. 4g), but this irregularity was found in a PDB (3O95\_B and 3O95\_D vs. 3O95\_A and 3O95\_C, 3O9V\_B and 3O9V\_D vs. 3O9V\_A and 3O9V\_C) despite having the same inhibitor. The reason that irregular orientations appeared for the Ser209 and Val711 side chains is unclear, although it is possible that the inhibitor influences their orientations.

Asn710 had 47 units in which the N $\delta$  and O $\delta$  atoms were switched to each other compared to the other inhibitor-bound units (Fig. 4h). In the commercial drugs, only linagliptin and alogliptin contained the switched Asn710. This trend was found in the inhibitor-free units (Additional file 3: Figure S2f). There were two fixed water O atoms found in both of the inhibitor-bound/free units, but their positions were not changed by switching the N $\delta$  and O $\delta$  atoms (Fig. 4h). No units had the inhibitor N atom close to (<4 Å) the switched/non-switched N $\delta$  atom of Asn710 (data not shown). This result suggests a tendency that the N $\delta$  and O $\delta$  atoms of Asn710 are switched to each other so that the N $\delta$  atom does not approach too close to the inhibitor's N atom.

The side chains of the above six residues may be flexible or have orientational alternatives according to the inhibitor used. However, the side chains of the other 10 residues (Arg125, Glu205, Glu206, Phe357, Tyr631, Val656, Trp659, Tyr662, Tyr666 and His740) had a less variation than the average of the equivalent amino acids (Fig. 3), and the superposition of the equivalent amino acids showed that their side chains had little orientational

difference between units (data not shown). These results suggest that the side chains of the 10 residues are mostly unaltered according to the inhibitor used. Little movement of the DPP-4 inhibitor binding area through the superposition of some X-ray cocrystal structures with different inhibitors has been described in some studies [35, 41, 45], and they are consistent with our suggestions. The observed rigidity is not a crystal property but probably a DPP-4 property because the collected structures were from different constructs and from different crystal forms, as many researchers have suggested.

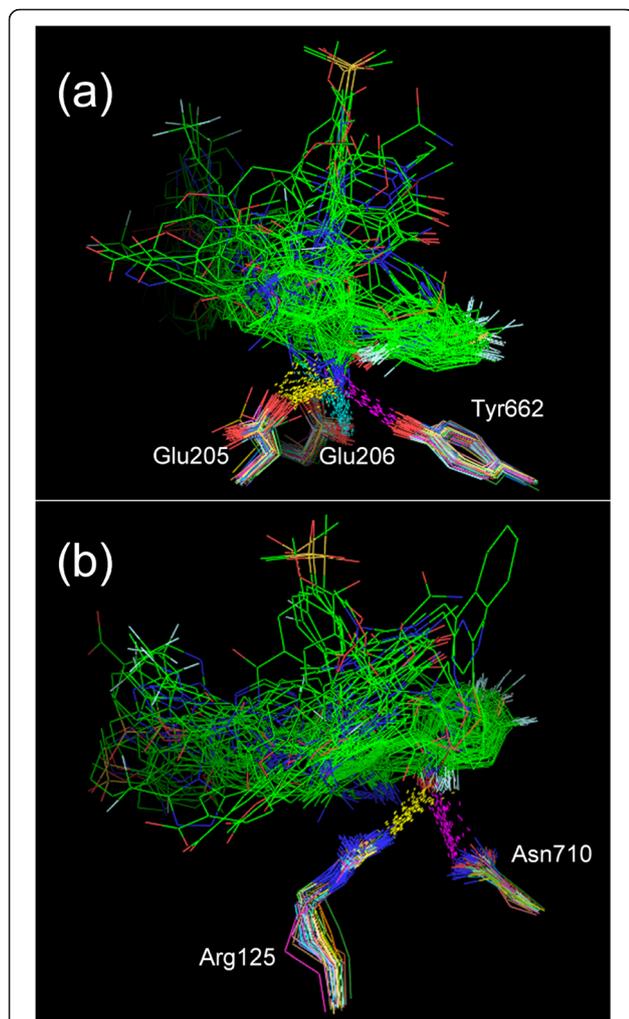
#### Common binding modes between DPP-4 and inhibitors

We searched electrostatic binding modes between DPP-4 and its inhibitors. In the DPP-4 inhibitor binding area, binding modes of Arg125, Glu205, Glu206, Tyr662 or Asn710 were common between many inhibitors.

Out of 68 inhibitors, 67 had a primary or secondary amino N atom close to one of two Glu205 O $\epsilon$  atoms (<4 Å, dashed yellow lines in Fig. 5a) and one or both of the two Glu206 O $\epsilon$  atoms (<4 Å, dashed cyan lines in Fig. 5a). The exception was 1TKR because the inhibitor in 1TKR has no N atom. In addition, the 67 inhibitors had the same N atom close to the Tyr662 O $\eta$  atom (<5 Å, dashed magenta lines in Fig. 5a). This network has been reported to be important for DPP-4 inhibitor to bind to the DPP-4 active center [2, 13–17, 20, 22, 30, 32–37, 40–42, 45–62]. The side chains of the three residues were oriented in a pre-determined direction, as discussed above. This means that a primary or secondary amino group of inhibitor is present without exception at a particular spatial position in DPP-4 and forms binding modes with the Glu205, Glu206 and Tyr662 side chains, suggesting that these binding modes may be a powerful rule for DPP-4 inhibitor to maintain stable binding with DPP-4.

Of 68 inhibitors, 53 had an O, N or halogen atom close to at least one of the Arg125 N $\eta$  atoms (Fig. 5b, dashed yellow lines, < 4 Å) and the Asn710 N $\delta$  or O $\delta$  atom (Fig. 5b, dashed magenta lines, < 4 Å). As mentioned above, the Arg125 and Asn710 side chains were oriented in a pre-determined direction, although position switching between the O $\delta$  and N $\delta$  atoms of Asn710 occurred in some units. The network of Asn710 and/or Arg125 also has been described by many authors [20, 32, 34, 35, 40, 41, 49, 52, 55, 56]. However, for the commercial drugs, only linagliptin (PDB ID: 2RGU [16]) had no binding mode with Arg125 and Asn710, and therefore this network may be less important than that with Glu205, Glu206 and Tyr662.

The Ser630 O $\gamma$  atom was also close to the polar atom of many inhibitors, but there were minimal common characteristics found on the inhibitor side. The disorderly orientation of the Ser630 O $\gamma$  atom was discussed



**Fig. 5** Superposition of DPP4 inhibitors highlights the important binding interactions to some residues. When the inhibitor-bound units were superimposed so that the RMSD targeting Ca atoms in DPP-4 (residue numbers 41–764) would be minimized based on 4PNZ\_A, 68 types of inhibitors were simultaneously superimposed. The carbon skeleton of inhibitors is indicated by green line. O, N and halogen atoms are colored red, blue and pale cyan, respectively. **a** Binding modes between Glu205, Glu206 or Tyr662 and inhibitors. The distances between a primary or secondary amino N atom of inhibitor and the Glu205 O $\epsilon$  atom ( $<4$  Å, yellow), the Glu206 O $\epsilon$  atom ( $<4$  Å, cyan), or the Tyr662 O $\eta$  atom ( $<5$  Å, magenta) are shown with dashed lines. **b** Binding modes between Arg125 or Asn710 and inhibitors. The distances between an O, N or halogen atom of inhibitor and the Arg125 N $\eta$  atom ( $<4$  Å, yellow) or the Asn710 O $\delta$ /N $\delta$  atom ( $<4$  Å, magenta) are shown with dashed lines

above, suggesting that the placement of an inhibitor with an excluded volume effect close to the Ser630 hydroxyl group is a main role for some DPP-4 inhibitors.

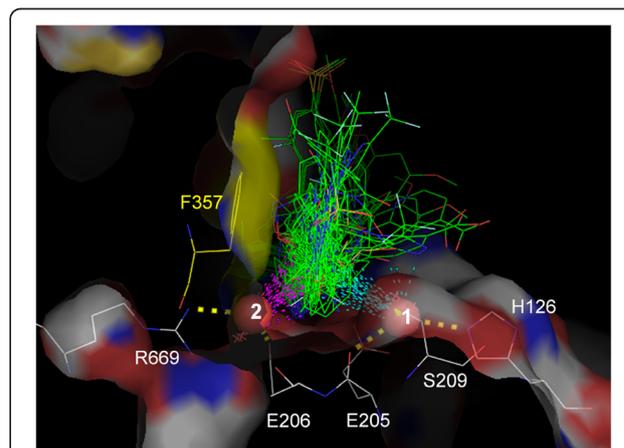
#### Common water molecules between units

When two units were superimposed, water O atoms within 1 Å were defined as identical. 2OQI\_B, 2OQV\_A,

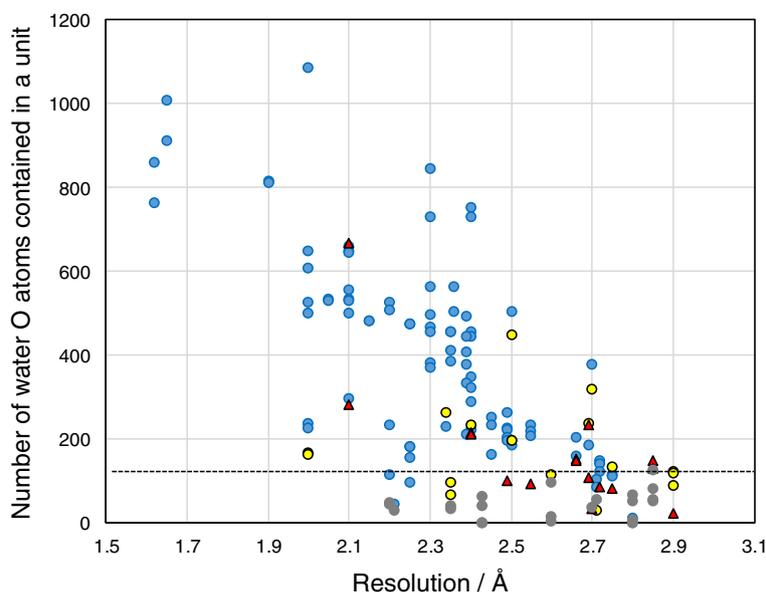
3F8S\_A and 3F8S\_B had no registered water O atoms. Two hundred thirty-eight water O atoms corresponded to the above identical definition in at least 61 out of the remaining 143 inhibitor-bound units. Among these, 26 water O atoms were close to the inhibitor binding area ( $<4$  Å), and two specific water O atoms were also close to many inhibitors ( $<5$  Å, red spheres in Fig. 6).

The first specific water O atom was observed in 108 inhibitor-bound units (blue circles and red triangles in Fig. 7). It was close to 52 inhibitors ( $<5$  Å, dashed cyan lines in Fig. 6), the His126 N $\epsilon$ , the Glu205 O $\epsilon$  and the Ser209 O $\gamma$  atoms ( $<4$  Å, dashed bold yellow lines in Fig. 6). This water molecule has been described in some studies [13, 45, 63]. The second specific water O atom was also observed in 108 units (blue and yellow circles in Fig. 7). It was close to 55 inhibitors ( $<5$  Å, dashed magenta lines in Fig. 6), the Glu206 O $\epsilon$  and the Arg669 N $\eta$  atoms ( $<4$  Å, dashed bold yellow lines in Fig. 6). Ninety-two units had both of the two water O atoms (blue circles in Fig. 7).

Many of the inhibitor-bound units that lack either of the two specific water O atoms were distributed over 2.5 Å resolution (red triangles, and yellow and grey circles in Fig. 7). Nineteen units that did not have the two O water atoms (grey circles in Fig. 7) showed few registered water O atoms (the maximum was 129 water O



**Fig. 6** Two specific water O atoms close to inhibitors (red spheres). Both of the two water O atoms were found in 108 inhibitor-bound units. When superimposed so that the RMSD targeting Ca atoms in DPP-4 (residue 41–764) would be minimized based on 1X70\_A, water O atoms in the unit were simultaneously superimposed. The carbon skeleton of inhibitors is shown with green line. O, N and halogen atoms are colored red, blue and pale cyan, respectively. The DPP-4 structure was obtained from the coordinates of PDB ID 1X70. The distances between inhibitors and the first specific water O atoms (No. 1) are shown with dashed cyan lines. The distances between inhibitors and the second specific water O atoms (No. 2) are shown with dashed magenta lines. The distances between the two water O atoms and His126, Glu205, Glu206, Ser209 or Arg669 are shown with dashed bold yellow line. Phe357 (yellow area) forms a large wall at the S2 subsite



**Fig. 7** Correlation between X-ray diffraction resolution and the number of water O atoms registered. Ninety-two units (blue circles) have both of the two specific water O atoms. Sixteen units (red triangles) have only the first specific water O atom (the first red sphere in Fig. 6), and 16 units (yellow circles) have only the second specific water O atom (the second red sphere in Fig. 6). Nineteen units (grey circles) have neither of the two water O atoms, and 2OQI\_B, 2OQV\_A, 3F8S\_A and 3F8S\_B (grey circles) have no registered water O atoms. The unit that contains the most water O atoms in the above 19 units (grey circles) is PDB ID 3O95\_B (129 water O atoms, dashed line)

atoms, dashed line in Fig. 7). These results suggest that the two specific water O atoms were overlooked in the units because their positions could not be identified accurately. Keedy et al. reports that the ordering of surface-associated waters is driven by cryocooling [31]. However, the high temperature-measured units had two water O atoms in the same position as the other inhibitor-bound units (Additional file 4: Figure S3), and the two specific water O atoms were also found in the inhibitor-free units (Additional file 5: Figure S4). The B-factor is an index of atom thermal mobility within crystal structure. In many inhibitor-bound units, both of the two specific water O atoms had a lower B-factor than the average B-factors of all water O atoms and of all protein heavy atoms (Additional file 6: Table S2 and Additional file 7: Table S3). Generally, hydrated water that is usually entrapped from local protein structure has a lower B-factor than crystal water that is accidentally entrapped by crystallization. The above two water molecules we found seem to exist as hydrated water that is taken into the local position in inherent DPP-4, independently of a temperature change (ranged from 90 to 298 K) and regardless if DPP-4 has an inhibitor.

The primary role of the two specific water molecules (probably hydrated water) may be forming an electrostatic network with His126, Glu205 and Ser209, and with Glu206 and Arg669, and maintaining the orientation of the side chains of Glu205 and Glu206 through the network may be especially important for inhibitor

binding. The adamantan hydroxyl group of vildagliptin is reported to interact with the first specific water molecule [13], and it may be important for the potency of vildagliptin. The two water molecules, however, do not seem to have an electrostatic interaction with many inhibitors because they are close to the carbon skeleton of the inhibitors rather than to their polar atoms. All 68 inhibitors were placed as if they were avoiding the position of the two water O atoms. The secondary role of the two water molecules may be an excluded volume effect, such as arranging inhibitor appropriately at the S2 subsite. The two water O atoms were located at the bottom of the S2 subsite, where Phe357 formed a large wall at the S2 subsite (yellow area in Fig. 6). The hydrophobic or  $\pi$ - $\pi$  stacking interaction of the Phe357 benzene ring has been reported to be important for binding in many inhibitors [2, 17, 22, 33, 42, 45, 47–49, 51–55, 57, 60, 64]. Further research is needed to determine the detailed relationship between the two specific water molecules and inhibitor potency. However, the pre-arrangement of the two water molecules should be considered in *in silico* screening not to lead inhibitor to improper position.

It has been often asked whether hydrated water in a crystal structure is applicable in solution. Generally, crystal and hydrated water occupies an average of 50 % of the crystal structure, and many protein crystal structures have been reported to reflect the structure in a water solution [65–77]. A molecular

dynamics simulation showed that the frequent appearance positions of water molecule during the simulation agrees with the hydrated water site in the X-ray crystal structure [78]. However, whether the two specific water molecules are necessary for an inhibitor to maintain stable binding with DPP-4 should be confirmed by a solution structure using nuclear magnetic resonance (NMR) or molecular dynamics simulation (the solution structure of DPP-4 has not yet been obtained using NMR). A few molecular dynamics simulations have reported the interaction between DPP-4 and its inhibitor [79, 80], but further simulations evaluating hydrated water require additional research, which we plan to perform in the future.

#### DPP-4 inhibitory affinity of the inhibitor used

The results so far were provided based on only crystal structures whose resolutions were less than 3.0 Å to provide high-quality information (Additional file 2: Table S1). However, if the results are used by *in silico* screening, a realistic potency should accompany the candidates that passed the screening test. DPP-4 inhibitory activity of PDBs (inhibitors) that we collected ranged from 0.1 to 30,000 nM in IC<sub>50</sub>/K<sub>i</sub>, and many of them had high inhibitory activity of less than 30 nM (Additional file 8: Table S4) [24, 25]. This value is a standard of the commercial drugs, as shown by Table 1. Thus, we used high-quality resources in regard to resolution and DPP-4 inhibitory affinity. Therefore, our study provides necessary information for *in silico* screening, and may provide a necessary minimum consensus to help in the discovery of a novel DPP-4 inhibitor that is commercially useful. However, DPP-4 inhibitor also needs to meet some optional requirements to acquire substantial commercial-drug level inhibitory activity, e.g. geometric fitness and electrostatic/non-electrostatic interaction with some subsites.

#### Conclusions

To clarify whether DPP-4 alters its general or partial structure according to the inhibitor used and whether DPP-4 has a common rule for inhibitor binding, we comprehensively analyzed X-ray cocrystal structures of DPP-4 and its inhibitor. Before beginning this research, we expected that DPP-4 would alter at least its partial structure because DPP-4 can bind to inhibitors of various shapes. However, our results were different from our initial expectations. All the main and side chains in the inhibitor binding area were only minimally altered except for some side chains (Ser209, Arg358, Tyr547, Ser630, Val656 and Asn710), despite binding to inhibitors of various shapes. Some residues (Arg125, Glu205, Glu206, Tyr662 and Asn710) in the area had binding modes to fix a specific atom of inhibitor to a particular

spatial position in DPP-4. We found two specific water molecules that are common to 92 DPP-4 crystal structures. The two water molecules were close to many inhibitors, and seemed to play two roles: maintaining the orientation of the Glu205 and Glu206 side chains, which is important for inhibitor binding through an electrostatic network via the two water molecules, and arranging the inhibitor appropriately at the S2 subsite. In this study, we used high-quality resources, and therefore this information may provide a necessary minimum consensus to help in the discovery of a novel DPP-4 inhibitor that is commercially useful.

#### Additional files

**Additional file 1: Figure S1.** Structural formula and PDB ID of 68 types of DPP-4 inhibitors. The PDB ID is followed by the chain ID that we used in this study. (DOCX 711 kb)

**Additional file 2: Table S1.** Resolution and measured temperature of X-ray crystal structural analysis. (DOCX 29 kb)

**Additional file 3: Figure S2.** Superposition of some residues in ligand-free units. Eight ligand-free units are superimposed so that the RMSD targeting the main chain (the N, Ca and C atoms) would be minimized based on a specific ligand-free unit. **a** Arg358, **b** Tyr547, **c** Ser630, **d** Ser209, **e** Val711 and **f** Asn710. (DOCX 124 kb)

**Additional file 4: Figure S3.** Superposition of two specific water O atoms in high temperature-measured units (yellow symbol "+"). The temperature-measured units were superimposed so that Ca atoms in DPP-4 (residue numbers 41–764) would be minimized based on 1X70\_A (the two specific water O atoms of 1X70\_A are marked by red symbol "+" in the white broken circles). The two specific water O atoms registered in the high temperature-measured units are superimposed well onto those in 1X70\_A that was measured at low temperature (yellow symbol "+" in the white broken circles). (DOCX 273 kb)

**Additional file 5: Figure S4.** Superposition of two specific water O atoms in ligand-free units (yellow spheres). Six ligand-free units (1J2E\_A, 1J2E\_B, 1NU6\_A, 1PFQ\_A, 1TK3\_A and 1TK3\_B) that had one of the two specific water O atoms were superimposed so that Ca atoms in DPP-4 (residue numbers 41–764) would be minimized based on 1X70\_A. The two specific water O atoms of the six ligand-free units (yellow spheres in the white broken circles) superimposed well onto those in 1X70\_A with ligand (red spheres). (DOCX 329 kb)

**Additional file 6: Table S2.** B-factor of the first specific water O atom in a unit. (DOCX 40 kb)

**Additional file 7: Table S3.** B-factor of the second specific water O atom in a unit. (DOCX 41 kb)

**Additional file 8: Table S4.** DPP-4 inhibitory activity. (DOCX 30 kb)

#### Abbreviations

DPP-4, dipeptidyl peptidase IV; GLP-1, glucagon-like peptide-1; IC<sub>50</sub>, half maximal inhibitory concentration; K<sub>i</sub>, inhibition constant; nM, nanomolar; NMR, nuclear magnetic resonance; PDB, Protein Data Bank; RMSD, root mean square deviation

#### Acknowledgements

We would like to thank Dr. Hiromitsu Shimoyama, Shogo Koike and Misa Obata for their support and helpful discussion.

#### Funding

Not applicable.

**Availability of data and materials**

All results are presented in the main text and the following additional supporting files provided in.docx format.

**Authors' contributions**

HN generated the original idea, executed the research and wrote the manuscript. KK, GT, MT-S, GI, KA, Clg and Clt conducted the research. HM supervised the project and edited the paper. All authors have read and approved the final manuscript.

**Competing interests**

The authors declare that they have no competing interests.

**Consent for publication**

Not applicable.

**Ethics approval and consent to participate**

Not applicable.

**Author details**

<sup>1</sup>School of Pharmacy, Kitasato University, 5-9-1 ShirokaneMinato-ku, Tokyo 108-8641, Japan. <sup>2</sup>Present address: Infectious Disease Surveillance Center, National Institute of Infectious Diseases, 1-23-1 ToyamaShinjuku-ku, Tokyo 162-8640, Japan.

Received: 2 November 2015 Accepted: 29 July 2016

Published online: 05 August 2016

**References**

- Zander M, Madsbad S, Madsen JL, Holst JJ. Effect of 6-week course of glucagon-like peptide 1 on glycaemic control, insulin sensitivity, and  $\beta$ -cell function in type 2 diabetes; a parallel-group study. *Lancet*. 2002;359:824–30.
- Kim D, Wang L, Beconi M, Eiermann GJ, Fisher MH, He H, et al. (2R)-4-oxo-4-[3-(trifluoromethyl)-5,6-dihydro[1,2,4]triazolo[4,3-a]pyrazin-7(8H)-yl]-1-(2,4,5-trifluorophenyl)butan-2-amine: a potent, orally active dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. *J Med Chem*. 2005;48(1):141–51.
- Nauck MA, Heimesaat MM, Orskov C, Holst JJ, Ebert R, Creutzfeldt W. Preserved incretin activity of glucagon-like peptide 1 [7–36 amide] but not of synthetic human gastric inhibitory polypeptide in patients with type-2 diabetes mellitus. *J Clin Invest*. 1993;91(1):301–7.
- Kim HM, Lim JS, Lee BW, Kang ES, Lee HC, Cha BS. Optimal candidates for the switch from glimepiride to sitagliptin to reduce hypoglycemia in patients with type 2 diabetes mellitus. *Endocrinol Metab (Seoul)*. 2015;30(1):84–91.
- Schmidt WE, Siegel EG, Creutzfeldt W. Glucagon-like peptide-1 but not glucagon-like peptide-2 stimulates insulin release from isolated rat pancreatic islets. *Diabetologia*. 1985;28(9):704–7.
- Shima K, Hirota M, Ohboshi C. Effect of glucagon-like peptide-1 on insulin secretion. *Regul Pept*. 1988;22(3):245–52.
- Ahrén B, Schmitz O. GLP-1 receptor agonists and DPP-4 inhibitors in the treatment of type 2 diabetes. *Horm Metab Res*. 2004;36(11–12):867–76.
- Hermansen K, Bækdal TA, Düring M, Pietraszek A, Mortensen LS, Jørgensen H, et al. Liraglutide suppresses postprandial triglyceride and apolipoprotein B48 elevations after a fat-rich meal in patients with type 2 diabetes: a randomized, double-blind, placebo-controlled, cross-over trial. *Diabetes Obes Metab*. 2013;15(11):1040–8.
- Kothare PA, Linnebjerg H, Isaka Y, Uenaka K, Yamamura A, Yeo KP, et al. Pharmacokinetics, pharmacodynamics, tolerability, and safety of exenatide in Japanese patients with type 2 diabetes mellitus. *J Clin Pharmacol*. 2008;48(12):1389–99.
- Petersen AB, Knop FK, Christensen M. Lixisenatide for the treatment of type 2 diabetes. *Drugs Today (Barc)*. 2013;49(9):537–53.
- Cunningham DF, O'Connor B. Proline specific peptidases. *Biochim Biophys Acta*. 1997;1343(2):160–86.
- Brandt I, Joossens J, Chen X, Maes MB, Scharpé S, De Meester I, et al. Inhibition of dipeptidyl-peptidase IV catalyzed peptide truncation by Vildagliptin ((2S)-{[(3-hydroxyadamantan-1-yl)amino]acetyl}-pyrrolidine-2-carbonitrile). *Biochem Pharmacol*. 2005;70(1):134–43.
- Nabeno M, Akahoshi F, Kishida H, Miyaguchi I, Tanaka Y, Ishii S, et al. 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):191–6.
- Feng J, Zhang Z, Wallace MB, Stafford JA, Kaldor SW, Kassel DB, et al. Discovery of alogliptin: a potent, selective, bioavailable, and efficacious inhibitor of dipeptidyl peptidase IV. *J Med Chem*. 2007;50(10):2297–300.
- Zhang Z, Wallace MB, Feng J, Stafford JA, Skene RJ, Shi L, et al. Design and synthesis of pyrimidinone and pyrimidinedione inhibitors of dipeptidyl peptidase IV. *J Med Chem*. 2011;54(2):510–24.
- Eckhardt M, Langkopf E, Mark M, Tadayyon M, Thomas L, Nar H, et al. 8-(3-(R)-aminopiperidin-1-yl)-7-but-2-ynyl-3-methyl-1-(4-methyl-quinazolin-2-ylmethyl)-3,7-dihydropurine-2,6-dione (BI 1356), a highly potent, selective, long-acting, and orally bioavailable DPP-4 inhibitor for the treatment of type 2 diabetes. *J Med Chem*. 2007;50(26):6450–3.
- Watanabe YS, Yasuda Y, Kojima Y, Okada S, Motoyama T, Takahashi R, et al. Anagliptin, a potent dipeptidyl peptidase IV inhibitor: its single-crystal structure and enzyme interactions. *J Enzyme Inhib Med Chem*. 2015;30(6):981–8.
- Kato N, Oka M, Murase T, Yoshida M, Sakairi M, Yamashita S, et al. Discovery and pharmacological characterization of N-[2-({(2S)-2-cyanopyrrolidin-1-yl}-2-oxoethyl)amino]-2-methylpropyl]-2-methylpyrazolo[1,5-a]pyrimidine-6-carboxamide hydrochloride (anagliptin hydrochloride salt) as a potent and selective DPP-IV inhibitor. *Bioorg Med Chem*. 2011;19(23):7221–7.
- Yoshida T, Akahoshi F, Sakashita H, Kitajima H, Nakamura M, Sonda S, et al. Discovery and preclinical profile of teneligliptin (3-((2S,4S)-4-[4-(3-methyl-1-phenyl-1H-pyrazol-5-yl)piperazin-1-yl]pyrrolidin-2-ylcarbonyl)thiazolidine): a highly potent, selective, long-lasting and orally active dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. *Bioorg Med Chem*. 2012;20(19):5705–19.
- Metzler WJ, Yanchunas J, Weigelt C, Kish K, Klei HE, Xie D, et al. Involvement of DPP-IV catalytic residues in enzyme-saxagliptin complex formation. *Protein Sci*. 2008;17(2):240–50.
- McKeage K. Trelagliptin: First global approval. *Drugs*. 2015;75(10):1161–4.
- Biftu T, Sinha-Roy R, Chen P, Qian X, Feng D, Kuethe JT, et al. Omarigliptin (MK-3102): a novel long-acting DPP-4 inhibitor for once-weekly treatment of type 2 diabetes. *J Med Chem*. 2014;57(8):3205–12.
- Kawabata T. Build-up algorithm for atomic correspondence between chemical structures. *J Chem Info Model*. 2011;51(8):1775–87. <http://strcomp.protein.osaka-u.ac.jp/kcombu/>. Accessed 28 May 2015.
- Chen X, Liu M, Gilson MK. BindingDB: a web-accessible molecular recognition database. *Comb Chem High Throughput Screen*. 2001;4(8):719–25. <http://www.bindingdb.org/bind/index.jsp>. Accessed 11 June 2016.
- Wang R, Fang X, Lu Y, Wang S. The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J Med Chem*. 2004;47(12):2977–80. <http://www.pdbbind.org.cn/>. Accessed 11 June 2016.
- Hiramatsu H, Kyono K, Higashiyama Y, Fukushima C, Shima H, Sugiyama S, et al. The structure and function of human dipeptidyl peptidase IV, possessing a unique eight-bladed beta-propeller fold. *Biochem Biophys Res Commun*. 2003;302(4):849–54.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. *Nucleic Acids Res*. 2000;28(1):235–42.
- Grell L, Parkin C, Slate L, Craig PA. EZ-Viz, A tool for simplifying molecular viewing in PyMOL. *Biochem Mol Biol Educ*. 2006;34(6):402–7.
- Fraczkiewicz R, Braun W. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J Comp Chem*. 1998;19:319–33. <http://curie.utmb.edu/getarea.html>. Accessed 15 January 2016.
- Sutton JM, Clark DE, Dunsdon SJ, Fenton G, Fillmore A, Harris NV, et al. Novel heterocyclic DPP-4 inhibitors for the treatment of type 2 diabetes. *Bioorg Med Chem Lett*. 2012;22(3):1464–8.
- Keedy DA, van den Bedem H, Sivak DA, Petsko GA, Ringe D, Wilson MA, et al. Crystal cryocooling distorts conformational heterogeneity in a model Michaelis complex of DHFR. *Structure*. 2014;22(6):899–910.
- Sheehan SM, Mest HJ, Watson BM, Klimkowski VJ, Timm DE, Cauvin A, et al. Discovery of non-covalent dipeptidyl peptidase IV inhibitors which induce a conformational change in the active site. *Bioorg Med Chem Lett*. 2007;17(6):1765–8.
- Pei Z, Li X, Longenecker K, von Geldern TW, Wiedeman PE, Lubben TH, et al. Discovery, structure-activity relationship, and pharmacological evaluation of (5-substituted-pyrrolidinyl-2-carbonyl)-2-cyanopyrrolidines as potent dipeptidyl peptidase IV inhibitors. *J Med Chem*. 2006;49(12):3520–35.
- Xu J, Wei L, Mathvink R, Edmondson SD, Mastracchio A, Eiermann GJ, et al. Discovery of potent, selective, and orally bioavailable pyridone-based dipeptidyl peptidase-4 inhibitors. *Bioorg Med Chem Lett*. 2006;16(5):1346–9.

35. Edmondson SD, Wei L, Xu J, Shang J, Xu S, Pang J, et al. Fluoroolefins as amide bond mimics in dipeptidyl peptidase IV inhibitors. *Bioorg Med Chem Lett.* 2008;18(7):2409–13.
36. Wallace MB, Feng J, Zhang Z, Skene RJ, Shi L, Caster CL, et al. Structure-based design and synthesis of benzimidazole derivatives as dipeptidyl peptidase IV inhibitors. *Bioorg Med Chem Lett.* 2008;18(7):2362–7.
37. Boehringer M, Fischer H, Hennig M, Hunziker D, Huwyler J, Kuhn B, et al. Aryl- and heteroaryl-substituted aminobenzo[a]quinolizines as dipeptidyl peptidase IV inhibitors. *Bioorg Med Chem Lett.* 2010;20(3):1106–8.
38. Wang W, Devasthale P, Wang A, Harrity T, Egan D, Morgan N, et al. 7-Oxopyrrolopyridine-derived DPP4 inhibitors-mitigation of CYP and hERG liabilities via introduction of polar functionalities in the active site. *Bioorg Med Chem Lett.* 2011;21(22):6646–51.
39. Lam B, Zhang Z, Stafford JA, Skene RJ, Shi L, Gwaltney 2nd SL. Structure-based design of pyridopyrimidinediones as dipeptidyl peptidase IV inhibitors. *Bioorg Med Chem Lett.* 2012;22(21):6628–31.
40. Edmondson SD, Mastracchio A, Mathvink RJ, He J, Harper B, Park YJ, et al. (2S,3S)-3-Amino-4-(3,3-difluoropyrrolidin-1-yl)-N, N-dimethyl-4-oxo-2-[4-[1,2,4]triazolo[1,5-a]pyridin-6-yl]phenyl]butanamide: a selective alpha-amino amide dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. *J Med Chem.* 2006;49(12):3614–27.
41. Liang GB, Qian X, Biftu T, Singh S, Gao YD, Scapin G, et al. Discovery of new binding elements in DPP-4 inhibition and their applications in novel DPP-4 inhibitor design. *Bioorg Med Chem Lett.* 2008;18(13):3706–10.
42. Qiao L, Baumann CA, Crysler CS, Ninan NS, Abad MC, Spurlino JC, et al. Discovery, SAR, and X-ray structure of novel biaryl-based dipeptidyl peptidase IV inhibitors. *Bioorg Med Chem Lett.* 2006;16(1):123–8.
43. Madar DJ, Kopecka H, Pireh D, Yong H, Pei Z, Li X, et al. Discovery of 2-[4-{{2-(2S,5R)-2-cyano-5-ethynyl-1-pyrrolidinyl}-2-oxoethyl]amino]-4-methyl-1-piperidinyl]-4-pyridinecarboxylic acid (ABT-279): a very potent, selective, effective, and well-tolerated inhibitor of dipeptidyl peptidase-IV, useful for the treatment of diabetes. *J Med Chem.* 2006;49(21):6416–20.
44. Asakura M, Fujii H, Atsuda K, Itoh T, Fujiwara R. Dipeptidyl peptidase-4 greatly contributes to the hydrolysis of vildagliptin in human liver. *Drug Metab Dispos.* 2015;43(4):477–84.
45. Biftu T, Scapin G, Singh S, Feng D, Becker JW, Eiermann G, et al. Rational design of a novel, potent, and orally bioavailable cyclohexylamine DPP-4 inhibitor by application of molecular modeling and X-ray crystallography of sitagliptin. *Bioorg Med Chem Lett.* 2007;17(12):3384–7.
46. Peters JU, Weber S, Ritter S, Weiss P, Wallier A, Boehringer M, et al. Aminomethylpyrimidines as novel DPP-IV inhibitors: a 10(5)-fold activity increase by optimization of aromatic substituents. *Bioorg Med Chem Lett.* 2004;14(6):1491–3.
47. Nordhoff S, Cerezo-Gálvez S, Feurer A, Hill O, Matassa VG, Metz G. The reversed binding of beta-phenethylamine inhibitors of DPP-IV: X-ray structures and properties of novel fragment and elaborated inhibitors. *Bioorg Med Chem Lett.* 2006;16(6):1744–8.
48. Pei Z, Li X, von Geldern TW, Madar DJ, Longenecker K, Yong H, et al. Discovery of ((4R,5S)-5-amino-4-(2,4,5-trifluorophenyl)cyclohex-1-enyl)-(3-(trifluoromethyl)-5,6-dihydro-[1,2,4]triazolo[4,3-a]pyrazin-7(8H)-yl)methanone (ABT-341), a highly potent, selective, orally efficacious, and safe dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. *J Med Chem.* 2006;49(22):6439–42.
49. Biftu T, Feng D, Qian X, Liang GB, Kieczkowski G, Eiermann G. (3R)-4-[(3R)-3-Amino-4-(2,4,5-trifluorophenyl)butanoyl]-3-(2,2,2-trifluoroethyl)-1,4-diazepan-2-one, a selective dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. *Bioorg Med Chem Lett.* 2007;17(1):49–52.
50. Lübbers T, Böhringer M, Gobbi L, Hennig M, Hunziker D, Kuhn B, et al. 1,3-disubstituted 4-aminopiperidines as useful tools in the optimization of the 2-aminobenzo[a]quinolizine dipeptidyl peptidase IV inhibitors. *Bioorg Med Chem Lett.* 2007;17(11):2966–70.
51. Backes BJ, Longenecker K, Hamilton GL, Stewart K, Lai C, Kopecka H, et al. Pyrrolidine-constrained phenethylamines: The design of potent, selective, and pharmacologically efficacious dipeptidyl peptidase IV (DPP4) inhibitors from a lead-like screening hit. *Bioorg Med Chem Lett.* 2007;17(7):2005–12.
52. Duffy JL, Kirk BA, Wang L, Eiermann GJ, He H, Leiting B, et al. 4-aminophenylalanine and 4-aminocyclohexylalanine derivatives as potent, selective, and orally bioavailable inhibitors of dipeptidyl peptidase IV. *Bioorg Med Chem Lett.* 2007;17(10):2879–85.
53. Pei Z, Li X, von Geldern TW, Longenecker K, Pireh D, Stewart KD, et al. Discovery and structure-activity relationships of piperidinone- and piperidine-constrained phenethylamines as novel, potent, and selective dipeptidyl peptidase IV inhibitors. *J Med Chem.* 2007;50(8):1983–7.
54. Wright SW, Ammirati MJ, Andrews KM, Brodeur AM, Danley DE, Doran SD, et al. (3R,4S)-4-(2,4,5-Trifluorophenyl)pyrrolidin-3-ylamine inhibitors of dipeptidyl peptidase IV: synthesis, in vitro, in vivo, and X-ray crystallographic characterization. *Bioorg Med Chem Lett.* 2007;17(20):5638–42.
55. Kowalchick JE, Leiting B, Pryor KD, Marsilio F, Wu JK, He H, et al. Design, synthesis, and biological evaluation of triazolo-piperazine-based beta-amino amides as potent, orally active dipeptidyl peptidase IV (DPP-4) inhibitors. *Bioorg Med Chem Lett.* 2007;17(21):5934–9.
56. Kaelin DE1, Smenton AL, Eiermann GJ, He H, Leiting B, Lyons KA, et al. 4-arylcyclohexylalanine analogs as potent, selective, and orally active inhibitors of dipeptidyl peptidase IV. *Bioorg Med Chem Lett.* 2007;17(21):5806–11.
57. Ammirati MJ, Andrews KM, Boyer DD, Brodeur AM, Danley DE, Doran SD, et al. (3,3-Difluoro-pyrrolidin-1-yl)-[(2S,4S)-(4-(4-pyrimidin-2-yl-piperazin-1-yl)-pyrrolidin-2-yl)]methanone: a potent, selective, orally active dipeptidyl peptidase IV inhibitor. *Bioorg Med Chem Lett.* 2009;19(7):1991–5.
58. Mattei P, Boehringer M, Di Giorgio P, Fischer H, Hennig M, Huwyler J, et al. Discovery of carmegliptin: a potent and long-acting dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. *Bioorg Med Chem Lett.* 2010;20(3):1109–13.
59. Miyamoto Y, Banno Y, Yamashita T, Fujimoto T, Oi S, Moritoh Y, et al. Discovery of a 3-pyridylacetic acid derivative (TAK-100) as a potent, selective and orally active dipeptidyl peptidase IV (DPP-4) inhibitor. *J Med Chem.* 2011;54(3):831–50.
60. Banno Y, Miyamoto Y, Sasaki M, Oi S, Asakawa T, Kataoka O, et al. Identification of 3-aminomethyl-1,2-dihydro-4-phenyl-1-isoquinolones: a new class of potent, selective, and orally active non-peptide dipeptidyl peptidase IV inhibitors that form a unique interaction with Lys554. *Bioorg Med Chem.* 2011;19(16):4953–70.
61. Devasthale P, Wang Y, Wang W, Fevig J, Feng J, Wang A, et al. Optimization of activity, selectivity, and liability profiles in 5-oxopyrrolopyridine DPP4 inhibitors leading to clinical candidate (Sa)-2-(3-(aminomethyl)-4-(2,4-dichlorophenyl)-2-methyl-5-oxo-5H-pyrrolo[3,4-b]pyridin-6(7H)-yl)-N,N-dimethylacetamide (BMS-767778). *J Med Chem.* 2013;56(18):7343–57.
62. Thoma R, Löffler B, Stihle M, Huber W, Ruf A, Hennig M. Structural basis of proline-specific exopeptidase activity as observed in human dipeptidyl peptidase-IV. *Structure.* 2003;11(8):947–59.
63. Ahn JH, Shin MS, Jun MA, Jung SH, Kang SK, Kim KR, et al. Synthesis, biological evaluation and structural determination of beta-aminoacyl-containing cyclic hydrazine derivatives as dipeptidyl peptidase IV (DPP-IV) inhibitors. *Bioorg Med Chem Lett.* 2007;17(9):2622–8.
64. Edmondson SD, Mastracchio A, Cox JM, Eiermann GJ, He H, Lyons KA, et al. Aminopiperidine-fused imidazoles as dipeptidyl peptidase-IV inhibitors. *Bioorg Med Chem Lett.* 2009;19(15):4097–101.
65. Blake PR, Day MW, Hsu BT, Joshua-Tor L, Park JB, Hare DR, et al. Comparison of the X-ray structure of native rubredoxin from *Pyrococcus furiosus* with the NMR structure of the zinc-substituted protein. *Protein Sci.* 1992;1:1522–5.
66. Braun W, Vasák M, Robbins AH, Stout CD, Wagner G, Kägi JH, et al. Comparison of the NMR solution structure and the x-ray crystal structure of rat metallothionein-2. *Proc Natl Acad Sci U S A.* 1992;89(21):10124–8.
67. Li A, Daggett V. Investigation of the solution structure of chymotrypsin inhibitor 2 using molecular dynamics: comparison to x-ray crystallographic and NMR data. *Protein Eng.* 1995;8(11):1117–28.
68. Blanco FJ, Ortiz AR, Serrano L. 1H and 15 N NMR assignment and solution structure of the SH3 domain of spectrin: comparison of unrefined and refined structure sets with the crystal structure. *J Biomol NMR.* 1997;9(4):347–57.
69. Sumikawa H, Suzuki E. Tertiary structural models of human interleukin-6 and evaluation by comparison with X-ray and NMR structures. *Chem Pharm Bull (Tokyo).* 1998;46(1):136–8.
70. Nicotra M, Paci M, Sette M, Oakley AJ, Parker MW, Lo Bello M, et al. Solution structure of glutathione bound to human glutathione transferase P1-1: comparison of NMR measurements with the crystal structure. *Biochemistry.* 1998;37(9):3020–7.
71. Gomar J, Sodano P, Sy D, Shin DH, Lee JY, Suh SW, et al. Comparison of solution and crystal structures of maize nonspecific lipid transfer protein: a model for a potential in vivo lipid carrier protein. *Proteins.* 1998;31(2):160–71.

72. Fraenkel E, Pabo CO. Comparison of X-ray and NMR structures for the Antennapedia homeodomain-DNA complex. *Nat Struct Biol.* 1998;5(8):692–7.
73. Lu J, Lin CL, Tang C, Ponder JW, Kao JL, Cistola DP, et al. The structure and dynamics of rat apo-cellular retinol-binding protein II in solution: comparison with the X-ray structure. *J Mol Biol.* 1999;286(4):1179–95.
74. Philippopoulos M, Lim C. Exploring the dynamic information content of a protein NMR structure: comparison of a molecular dynamics simulation with the NMR and X-ray structures of Escherichia coli ribonuclease HI. *Proteins.* 1999;36(1):87–110.
75. Haliloglu T, Bahar I. Structure-based analysis of protein dynamics: comparison of theoretical results for hen lysozyme with X-ray diffraction and NMR relaxation data. *Proteins.* 1999;37(4):654–67.
76. Kuser PR, Franzoni L, Ferrari E, Spisni A, Polikarpov I. The X-ray structure of a recombinant major urinary protein at 1.75 Å resolution. A comparative study of X-ray and NMR-derived structures. *Acta Crystallogr D Biol Crystallogr.* 2001;57(Pt 12):1863–9.
77. Kowalski JA, Liu K, Kelly JW. NMR solution structure of the isolated Apo Pin1 WW domain: comparison to the x-ray crystal structures of Pin1. *Biopolymers.* 2002;63(2):111–21.
78. Higo J, Nakasako M. Hydration structure of human lysozyme investigated by molecular dynamics simulation and cryogenic X-ray crystal structure analyses: on the correlation between crystal water sites, solvent density, and solvent dipole. *J Comput Chem.* 2002;23(14):1323–36.
79. Chakraborty C, Hsu MJ, Agoramoorthy G. Understanding the molecular dynamics of type-2 diabetes drug target DPP-4 and its interaction with sitagliptin and inhibitor diprotin-A. *Cell Biochem Biophys.* 2014;70(2):907–22.
80. Gu Y, Wang W, Zhu X, Dong K. Molecular dynamic simulations reveal the mechanism of binding between xanthine inhibitors and DPP-4. *J Mol Model.* 2014;20(2):2075.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

