Dipeptidyl peptidase-IV inhibitors used in type-2 diabetes inhibit a phospholipase C: a case of promiscuous scaffolds in proteins
[version 1; referees: 1 approved, 1 approved with reservations]

Sandeep Chakraborty¹,²*, Adela Rendón-Ramírez³*, Bjarni Ásgeirsson⁴*, Mouparna Dutta⁵, Anindya S. Ghosh⁵, Masataka Oda⁶, Ravindra Venkatramani⁷, Basuthkar J. Rao¹, Abhaya M. Dandekar², Félix M. Goñi³

¹Department of Biological Sciences, Tata Institute of Fundamental Research, Mumbai, 400 005, India
²Plant Sciences Department, University of California, Davis, CA, 95616, USA
³Unidad de Bio, Universidad del País Vasco, Bilbao, Spain
⁴Science Institute, Department of Biochemistry, University of Iceland, IS-107 Reykjavik, Iceland
⁵Department of Biotechnology, Indian Institute of Technology Kharagpur, Kharagpur, 721302, India
⁶Division of Microbiology and Infectious Diseases, Niigata University Graduate School of Medical and Dental Sciences, Niigata, 951-8514, Japan
⁷Department of Chemical Sciences, Tata Institute of Fundamental Research, Mumbai, 400 005, India

* Equal contributors

Abstract

The long term side effects of any newly introduced drug is a subject of intense research, and often raging controversies. One such example is the dipeptidyl peptidase-IV (DPP4) inhibitor used for treating type 2 diabetes, which is inconclusively implicated in increased susceptibility to acute pancreatitis. Previously, based on a computational analysis of the spatial and electrostatic properties of active site residues, we have demonstrated that phosphoinositide-specific phospholipase C (PI-PLC) from Bacillus cereus is a prolyl peptidase using in vivo experiments. In the current work, we first report the inhibition of the native activity of PI-PLC by two DPP4 inhibitors - vildagliptin (LAF-237) and K-579. While vildagliptin inhibited PI-PLC at micromolar concentrations, K-579 was a potent inhibitor even at nanomolar concentrations. Subsequently, we queried a comprehensive, non-redundant set of 5000 human proteins (50% similarity cutoff) with known structures using serine protease (SPASE) motifs derived from trypsin and DPP4. A pancreatic lipase and a gastric lipase are among the proteins that are identified as proteins having promiscuous SPASE scaffolds that could interact with DPP4 inhibitors. The presence of such scaffolds in human lipases is expected since they share the same catalytic mechanism with PI-PLC. However our methodology also detects other proteins, often with a completely different enzymatic mechanism, that have significantly congruent domains with the SPASE motifs. The reported elevated levels of serum lipase, although contested, could be rationalized by inhibition of lipases reported here. Also, the methodology presented here can be easily adopted for other drugs, and provide the first line of filtering in the
identification of pathways that might be inadvertently affected due to promiscuous scaffolds in proteins.
**Introduction**

Oral glucose elicits a greater insulin response than intravenous glucose infusion, a phenomenon known as the incretin effect. This effect is mostly attributed to the intestinally derived hormones glucagon-like peptide-1 (GLP-1) and gastric inhibitory polypeptide (GIP). These hormones have a very short half-life as they are rapidly inactivated by the ubiquitous enzyme dipeptidyl peptidase-IV (DPP4). The finding that the incretin effect is impaired in subjects with type 2 diabetes led to two major types of GLP-1 based therapies - intravenously or sub-cutaneously administered GLP-1 mimetics that are resistant to DPP4 (exenatide, liraglutide, etc.), and the orally administered gliptins that prolong the physiological actions of incretin hormones by inhibiting DPP4 (sitagliptin, vildagliptin, etc.). Due to the multifarious roles played by the DPP4 enzyme, the possible side effects of these drugs (acute pancreatitis, pancreatic cancer, etc.) are strongly contested by researchers who argue that current statistics are insufficient to conclusively attribute these side effects to the otherwise beneficial GLP-1 drugs. Compound promiscuity is another phenomenon that might play a crucial role in determining the side effects of these therapies, although this aspect has rarely been pursued intensively.

Previous work by our group has established the spatial and electrostatic congruence in cognate residue pairs of the active site in proteins with the same functionality (CLASP). CLASP analysis indicated that the phosphoinositide-specific phospholipase C (PI-PLC) from Bacillus cereus has spatial and electrostatic congruence with a serine protease inhibitor AEBSF (IC$_{50}$ = 0.018 mM). The specificity of the protease activity was for a proline in the amino terminal, suggesting that PI-PLC is a prolyl peptidase, similar to the DPP4 enzyme. This finding led us to believe that the gliptins would have similar inhibitory effect on PI-PLC. In the current work, we have confirmed the inhibition of the native phospholipase activity of PI-PLC using two gliptins - vildagliptin (at µ-molar concentrations) and K579 (at nano-molar concentrations).

Subsequently, we used a motif derived from a DPP4 protein, in addition to the trypsin motif used previously, to query a comprehensive and non-redundant (50% sequence identity) list of ~5000 human proteins with known structures using CLASP, intending to identify other proteins that might be inhibited by the gliptins. From the set of proteins with significant congruent matches with these two motifs, we identified a pancreatic lipase and a gastric lipase, keeping the context of lipases, acute pancreatitis and GLP-1 based therapies in mind. Our findings rationalize the elevated levels of serum lipase found in patients undergoing DPP4 inhibitor based therapies, although these reports are in disagreement with other findings.

While it is logical and expected to find scaffolds that are congruent to trypsin and DPP4 active sites in lipases based on the current results and our previous findings, we also show the presence of the serine catalytic triad in close proximity to the active site residues of proteins which have a completely different enzymatic mechanism (for example, in glutaminyl cyclase which is a transferase). This corroborates the current belief that convergent evolution occurs more frequently than previously believed. Thus, we propose a rational method to identify proteins that might have unintended and undesirables interactions with newly introduced compounds, and substantiate our claims by demonstrating the inhibition of the native phospholipase activity of PI-PLC from B. cereus using gliptins that are used in type 2 diabetes therapy.

**Results**

**The active site motifs**

The active sites of serine proteases differ in their specificities owing to residues other than the conserved catalytic triad. Thus, in addition to the trypsin motif used previously (Asp102, Ser195 and His57 - PDBid 1A0J (Motif1), we choose another motif from a DPP4 enzyme (Asp708, Ser630 and His740 - PDBid:1N1M) (Motif2) (Table 1). Apart from the catalytic triad, we chose another

<table>
<thead>
<tr>
<th>PDB</th>
<th>Active site atoms (a,b,c,d)</th>
<th>ab</th>
<th>ac</th>
<th>ad</th>
<th>bc</th>
<th>bd</th>
<th>cd</th>
<th>Rmsd1</th>
<th>Rmsd2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRYPSIN(1A0J)</td>
<td>D102,S195 H57_A56</td>
<td>7.8</td>
<td>5.6</td>
<td>2.9</td>
<td>3.3</td>
<td>9.0</td>
<td>6.9</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>DPP4(1N1M)</td>
<td>D708,S630 H174,V711</td>
<td>7.6</td>
<td>5.4</td>
<td>2.6</td>
<td>2.6</td>
<td>5.6</td>
<td>3.4</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>PI-PLC (1PTD)</td>
<td>D67,S234 H32,168</td>
<td>8.2</td>
<td>6.2</td>
<td>4.1</td>
<td>4.1</td>
<td>11.5</td>
<td>9.2</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>PLASE (2OXE)</td>
<td>D195,S42 S282,G235</td>
<td>7.7</td>
<td>6.4</td>
<td>4.1</td>
<td>3.8</td>
<td>15.5</td>
<td>18.2</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>GPASE(1HLG)</td>
<td>D324,S153 H353,L326</td>
<td>7.5</td>
<td>5.0</td>
<td>2.9</td>
<td>3.8</td>
<td>6.7</td>
<td>5.8</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>GPASE(1HLG)</td>
<td>D324,S153 H353,L326</td>
<td>7.5</td>
<td>5.0</td>
<td>2.9</td>
<td>3.8</td>
<td>6.7</td>
<td>5.8</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>QC(3PB4)</td>
<td>D170,S187 H168,G224</td>
<td>7.5</td>
<td>4.8</td>
<td>3.4</td>
<td>3.3</td>
<td>10.7</td>
<td>8.0</td>
<td>0.4</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1.** Potential and spatial congruence of the active site residues in proteins queried using two motifs - Motif1 from Trypsin and Motif2 from DPP4. Rmsd1 and Rmsd2 are the root mean square deviation of the scaffold with respect to Motif1 and Motif2. DPP4 - dipeptidyl peptidase-IV, PI-PLC - phosphoinositide-specific phospholipase C, PLASE - human pancreatic lipase-related Protein Z, GPASE - human gastric lipase, QC - glutaminyl cyclase. D = Pairwise distance in Å. PD = Pairwise potential.
non-polar residue in order to increase the specificity of the matches
(Ala56 in Motif1 and Val711 in Motif2). This fourth residue is chosen
as the closest residue to any one of the catalytic triad residues.
Using the ability of CLASP to include stereochemically equivalent
residues, this last residue could be matched by another non-polar
residue - one of Gly, Ala, Val, Leu, Ile or Met. Further, it has been
seen that the second (ac) and fifth (bd) (Table 1) pairwise electro-
static potential differences (EPD) are not discriminatory - thus, this
pair is not used to score the EPD difference (although it is included
in the distance deviation score).

Inhibition of phosphoinositide-specific phospholipase C
(PI-PLC) using dipeptidyl peptidase-IV (DPP4) inhibitors
DPP4 (EC 3.4.14.5), a serine protease that is expressed in many
tissues (kidney, liver, lung, intestinal membranes, lymphocytes and
endothelial cells), cleaves peptides with Pro or Ala residues in the
second amino terminal position. Previously, we have experimentally
demonstrated the existence of the serine protease domain in PI-
PLC from Bacillus cereus - both by virtue of its proteolytic activity,
and the inhibition of its native activity on phospholipids in the
presence of serine protease inhibitors\textsuperscript{22}. Furthermore, the specificity
of the proteolytic activity indicated that it was a prolyl pepti-
dase - thus, leading us to believe that DPP4 inhibitors should have
a similar inhibitory effect on the PI-PLC enzyme. Table 1 shows
the presence of a congruent motif in the PI-PLC protein with both
Motif1 and Motif2. His32 and Asp67 are known to be a part of the
active site scaffold in PI-PLC\textsuperscript{32}. These proteins have completely dif-
ferent folds, and thus a superimposition (using both MUSTANG\textsuperscript{34}
and DECAAF\textsuperscript{35}) does not show any detectable similarity in their
structures (Supplementary Figure 1). Figure 1 shows the active
sites of these proteins, and the superimposition of these proteins
based on their catalytic residues\textsuperscript{36}. It can be seen that the closest
nonpolar residue to the catalytic triad in trypsin and PI-PLC (Ala56
in PDBid:1A0J, Ile68 in PDBid:1PTD) is differently placed from
Val711 in DPP4 (PDBid:1N1M). This is also indicated by the greater
RMSD (root mean square deviation) of the scaffold in PI-PLC to
Motif2 as compared to Motif1. The differences in the position of
peripheral residues is the source of the diverse specificities exhib-
ited by these proteases. Figure 2 shows the inhibition of PI-PLC
using two gliptins - vildagliptin(LAF-237)\textsuperscript{23} and K579\textsuperscript{24}. PI-PLC
converts hydrolysis of phospholipids to yield diacylglycerol and
a phosphoryl alcohol. In the absence of inhibitors enzyme addi-
tion to the vesicle suspension causes an increase in turbidity due to
vesicle aggregation (Figure 2a,c). Aggregation in turn occurs as a
result of formation of the enzyme end-product diacylglycerol\textsuperscript{36,37}. A
steady-state is reached under our conditions after 6–8 min. Addition
of either LAF-237 (vildagliptin) or K579 leads to an obvious inhibi-
tion of the enzyme activity. Dose-response curves for the inhibitors

Figure 1. The active site residues in Trypsin, DPP4 and PI-PLC. (a) Trypsin (PDBid:1A0J) (b) DPP4 (PDBid:1N1M) (c) PI-PLC (PDBid:1PTD)
(d) Superimposing the active site residues using DE-CAAF\textsuperscript{36}.

Figure 2. The inhibition of PI-PLC with two gliptins.
are shown in Figure 2 (b,d). K579 is two orders of magnitude more potent than LAF-237 as a PI-PLC inhibitor, with half-maximal inhibitory concentrations IC\(_{50}\) respectively of 1 \(\mu\text{M}\) and 100 \(\mu\text{M}\).

Figure 2. PI-PLC inhibition using DPP4 inhibitors. (a,c) Time courses of enzyme activity in the presence of varying amounts of inhibitors, respectively LAF-237 and K579. The trace marked LIPOSOMES corresponds to a control in the absence of PI-PLC. (b,d) Dose-response effect of inhibitors on PI-PLC activity. Activity was computed as the extent of vesicle aggregation after 10 min enzyme activity.

Phosphoinositide-specific phospholipase C inhibition data using the dipeptidyl peptidase-IV inhibitors K-579 and LAF-237

12 Data Files

http://dx.doi.org/10.6084/m9.figshare.880620

Querying a non-redundant set of human proteins using Motif1 and Motif2

Currently, the PDB database has about 25,000 human proteins. Using an identity cutoff of 50\%, we chose a set of \(~5000\) proteins (Supplementary Table 1) as the target proteins. Table 2 shows the best matches obtained in these \(~5000\) proteins when queried by Motif1 and Motif2. Given the context of lipases, acute pancreatitis and GLP-1 based therapies, we picked two proteins - the human pancreatic lipase-related protein 2 (PDBid:2OXE) and a human gastric lipase (PDBid:1HLG) - to demonstrate the distinct possibility that these proteins might be inhibited by DPP4 inhibitors. Table 1 shows the congruence of the DPP4 motif to these proteins using Motif1 and Motif2. It is interesting to note that the gastric lipase (PDBid:1HLG) has a good match with both motifs - Leu326 in PDBid:1HLG is congruent to Ala56 in PDBid:1A0J, and Ala237 (PDBid:1HLG) is congruent to Val711 (PDBid:1N1M).

Since both these proteins are lipases (hydrolases), this congruence to Motif1 and Motif2 is expected based on our previous results with PI-PLC. However, our methodology also detects other proteins, often with a completely different enzymatic mechanism from hydrolases. A glutaminyl cyclase (PDBid:3PB4), a transferase, has a
Table 2. Best matches in the set of ∼5000 human proteins. (a) Motif1 (Asp102, Ser195, His57, Ala56) from Trypsin (b) Motif2 (Asp708, Ser630, His740, Val711) from DPP4.

<table>
<thead>
<tr>
<th>Motif</th>
<th>PDB</th>
<th>Description</th>
<th>CLASP Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2ANY</td>
<td>Plasma kallikrein, light chain</td>
<td>0.028</td>
</tr>
<tr>
<td>1</td>
<td>2OQ5</td>
<td>Transmembrane protease, serine 11E</td>
<td>0.037</td>
</tr>
<tr>
<td>1</td>
<td>3UOV</td>
<td>Lyso phospholipase-like protein 1</td>
<td>0.041</td>
</tr>
<tr>
<td>1</td>
<td>2ODP</td>
<td>Complement C2</td>
<td>0.060</td>
</tr>
<tr>
<td>1</td>
<td>1MJ</td>
<td>CCG1-interacting factor B</td>
<td>0.065</td>
</tr>
<tr>
<td>1</td>
<td>3F6U</td>
<td>Vitamin K-dependent protein C heavy chain</td>
<td>0.065</td>
</tr>
<tr>
<td>1</td>
<td>1ELV</td>
<td>Complement C1S component</td>
<td>0.068</td>
</tr>
<tr>
<td>1</td>
<td>1MD8</td>
<td>C1R complement serine protease</td>
<td>0.068</td>
</tr>
<tr>
<td>1</td>
<td>1ORF</td>
<td>Granzyme A</td>
<td>0.070</td>
</tr>
<tr>
<td>1</td>
<td>1FJ2</td>
<td>Acyl protein thioesterase 1</td>
<td>0.071</td>
</tr>
<tr>
<td>2</td>
<td>1HLG</td>
<td>Gastric lipase</td>
<td>0.042</td>
</tr>
<tr>
<td>2</td>
<td>1SPJ</td>
<td>Kallikrein 1</td>
<td>0.114</td>
</tr>
<tr>
<td>2</td>
<td>2F83</td>
<td>Coagulation factor XI</td>
<td>0.120</td>
</tr>
<tr>
<td>2</td>
<td>1ZJK</td>
<td>Mannan-binding lectin serine protease 2</td>
<td>0.131</td>
</tr>
<tr>
<td>2</td>
<td>3QLP</td>
<td>Thrombin light chain</td>
<td>0.145</td>
</tr>
<tr>
<td>2</td>
<td>2QXI</td>
<td>Kallikrein-7</td>
<td>0.146</td>
</tr>
<tr>
<td>2</td>
<td>2XU7</td>
<td>Histone-binding protein RBBP4</td>
<td>0.174</td>
</tr>
<tr>
<td>2</td>
<td>2W2N</td>
<td>Proprotein convertase subtilisin/kexin type 9</td>
<td>0.180</td>
</tr>
<tr>
<td>2</td>
<td>2HEH</td>
<td>KIF2C protein</td>
<td>0.195</td>
</tr>
<tr>
<td>2</td>
<td>2ANY</td>
<td>Plasma kallikrein, light chain</td>
<td>0.197</td>
</tr>
</tbody>
</table>

significantly congruent domain with Motif1 (lesser congruence with Motif2, as indicated by the RMSD) (Table 1). Figure 3 shows the proximity of the promiscuous scaffold to the active site of the cyclase, and also the congruence of the scaffold to Motif1.

Discussion
The controversy regarding the side effects of the dpp4 inhibitors, particularly with respect to acute pancreatitis and pancreatic cancer, continues unabated. While some researchers feel that it is not acceptable to assume that ‘absence of evidence is evidence of absence’38,39, others believe that current data are not conclusive and the ‘benefits by far outweigh the potential risks’16. Adding to the uncertainties are conflicting reports presented by different groups28–31. Notwithstanding the antagonistic views on the subject, it is unanimously accepted that current data are insufficient to establish a causal pathogenic effect of these drugs on such side effects40.

Various database studies have been undertaken in order to ascertain the effects of the GLP-1 therapies. Some studies ‘did not find an association between the use of exenatide or sitagliptin and acute pancreatitis’ with the caveat that the ‘limitations of this observational claims-based analysis cannot exclude the possibility of an increased risk’41. On the other hand, other studies have shown that the use of ‘sitagliptin or exenatide increased the odds ratio for reported pancreatitis 6-fold as compared with other therapies’14. Further, they reported that ‘pancreatic cancer was more commonly reported among patients who took sitagliptin or exenatide as compared with other therapies’14. The close relationship between chronic pancreatitis and pancreatic cancer is also a subject of intense research42. Another administrative database study of US adults with type 2 diabetes reported increased odds of hospitalization for acute pancreatitis for patients undergoing GLP-1-based therapies sitagliptin and exenatide14. Once again, such correlation of GLP-1 based therapies to acute pancreatitis is contested by other studies33.

Our findings rationalize the elevated levels of serum lipase found in patients undergoing DPP4 inhibitor based therapies42,29, keeping in mind that other studies contradict these reports30–31. While several studies have reported that the GLP-1 mimetics do not induce pancreatitis in rats, mouse and/or monkey44–46, these studies did not include DPP4 inhibitors, which are the compounds that might be responsible for interactions with pancreatic proteins according to our study. It is to be noted however that these mimetics may have other physiological effects and ‘the long-term consequences of sustained GLP-1 receptor activation in the human thyroid remain unknown and merit further investigation’47. Once again, the previous study47 has been challenged by another group who note that ‘findings previously reported in rodents may not apply to humans’48.
The orally administered gliptins differ in many aspects such as potency, excretion mechanism, target selectivity, half-life, metabolism and possible drug-drug interactions. This difference is also highlighted in the different concentrations of vildagliptin and K579 that inhibit PI-PLC. Interestingly, the PI-PLC scaffold has a better match with the trypsin motif than with the DPP4 motif (Table 1). In order to be able to model these differences in our in silico search, it is important to be able to provide flexibility in the scoring mechanism.

To summarize, it has been noted in the case of GLP-1 based therapies that as ‘evidence of harm accumulates, but is vigorously discounted’ the ‘burden of proof now rests with those who wish to convince us of their safety’. Surveillance programs, real-life cohort studies and case-control studies can be supplemented by rational investigations of relevant proteins based on anecdotal reports. The methodology proposed in the current work, which specifically demonstrates the effects of the DPP4 inhibitors, also presents a rational way of determining the inadvertent interactions of newly designed compounds with proteins, and thus prevent the recurrence of drug induced diseases being detected after considerable damage has already been inflicted on humans subjected to these drugs.

Materials and methods

In silico analysis

A comprehensive, non-redundant set of ~5000 human proteins (50% identity cutoff) was obtained from the PDB database. The CLASP package (http://www.sanchak.com/clasp) used for querying these proteins using motifs from trypsin and DPP4 is written in Perl on Ubuntu. Hardware requirements are modest - all results here are from a simple workstation (8GB ram), and runtimes for analyzing the ~5000 proteins was about 24 hours. Adaptive Poisson-Boltzmann Solver (APBS) and PDB2PQR packages were used to calculate the potential difference between the reactive atoms of the corresponding proteins. The APBS parameters and electrostatic potential units were set as described previously in Chakraborty et al. All protein structures were rendered by PyMol (http://www.py-mol.org/). Protein structures have been superimposed using MUS-TANG and DECAAF.

Protein, substrate and reagents

PI-PLC was purchased from Sigma. Vildagliptin(LAF-237) was obtained from Selleckchem, and K579 was obtained from Santa Cruz.

PI-PLC assay and inhibition using DPP4 inhibitors

Vesicle preparation and characterization. The appropriate lipids were mixed in organic solution, and the solvent was evaporated to dryness under N2. Solvent traces were removed by evacuating the lipids for at least 2 hours. The lipids were then swollen in 10 mM Hepes, 150 mM NaCl, pH 7.5 buffer. Large unilamellar vesicles (LUV) were prepared from the swollen lipids by extrusion and sized by using 0.1 μm poresize Nuclepore filters, as described by Ahyayauch et al. LUV composition was egg phosphatidylcholine: egg phosphatidylethanolamine: cholesterol at a 2:1:1 mole ratio. The average size of LUV was measured by quasi-elastic light scattering, using a Malvern Zeta-sizer instrument. Lipid concentration, determined by phosphate analysis, was 0.3 mM in all experiments.

Aggregation assay. Enzyme activity was assayed measuring enzyme-induced vesicle aggregation. All assays were carried out at 39°C with continuous stirring, in 10 mM Hepes, 150 mM NaCl buffer (pH 7.5), in the presence of 0.1% BSA for optimum catalytic activity. Enzyme concentration was 0.16 U/mL, and liposomal concentration was 0.3 mM. Lipid aggregation was monitored in a Cary Varian UV-vesicle spectrometer as an increase in turbidity (absorbance at 450 nm) of the sample, as described by Villar et al.
Author contributions
SC, ARR and BA performed the experiments. All authors analyzed the data, and contributed equally to the writing and subsequent refinement of the manuscript.

Competing interests
No competing interests were disclosed.

Grant information
FMG thanks the Spanish Ministerio de Ciencia e Innovacion for grant No. BFU 2012-36241, and the University of the Basque Country for grant No. IT 849-13. BJ and RV acknowledge financial support from Tata Institute of Fundamental Research (Department of Atomic Energy). Additionally, BJR is thankful to the Department of Science and Technology for the JC Bose Award Grant. BA extends gratitude to the University of Iceland Research Found for supporting the project financially. AMD wishes to acknowledge grant #12-0130-SA from California Department of Food and Agriculture CDFA PD/GWSS Board. MO was supported in part by a Grant-in-aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan; (grant No. 21790431).

Acknowledgements
We thank JM Frere for critical technical inputs at various stages of the project.

Supplementary material

Figure S1. Superimposition of trypsin (PDBid:1A0J - magenta), dipeptidyl peptidase-IV (PDBid:1N1M - yellow) and phosphoinositide-specific phospholipase C (PDBid:1PTD - cyan). It is seen that there is no structural similarity in the two proteins. (a) Using MUSTANG<sup>34</sup>. (b) Using DECAAF<sup>35</sup>.
Table S1.  PDZ IDs of ~5000 human proteins analyzed in this study.

<table>
<thead>
<tr>
<th>ID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A102</td>
<td>Table S1. PDB IDs of ~5000 human proteins analyzed in this study.</td>
</tr>
</tbody>
</table>


The successful targeting of DPP4 using small molecule compounds to treat type 2 diabetes has attracted a great deal of attention towards the study of this protease.

The authors applied sophisticated techniques that they have developed in order to discover that two DPP4 inhibitors, including one that is in limited clinical use, can to some extent inhibit the activity of a bacterial lipase (PI-PLC). Many lipases and esterases and hydrolases including DPP4 and related enzymes use the \textit{alpha/beta} hydrolase fold and the authors show how this related protein topology can place the residues in positions that are sufficiently similar to interact with an inhibitor.

The major difficulty with this paper is that it attempts to connect these data with possible clinical outcomes. No evidence for such a link is presented. Therefore, the title and much of the conclusions need to be modified so that they reflect the data without speculation.

Two inhibitors of DPP4, LAF237 and K-579, were studied. K-579 is not in clinical use. LAF237 is licensed in Europe and is known to exhibit some inhibition of the DPP4-related proteases DPP8 and DPP9. The extent of inhibition of DPP8 and DPP9 by LAF237 is believed not to have physiological effects in humans. The IC50 of LAF237 on DPP9 is less than 0.01 mM. The IC50 of LAF237 on bacterial PI-PLC is 0.1 mM, which is close to the lower limit of detection of inhibition of an enzyme. No mammalian homolog of PI-PLC was examined.

The literature that the authors cite to suggest that DPP4 inhibition might be detrimental for human health, particularly the pancreas, is data on sitagliptin or exenatide. Exenatide is not a DPP4 inhibitor and sitagliptin is quite different to LAF237, both in protease specificity and in chemical structure. The contact points of LAF237 and sitagliptin in the catalytic site of DPP4 differ considerably. The authors present no data on sitagliptin or any other DPP4 inhibitor (other than LAF237) that is in the clinic.

The images of overlaid catalytic triads of various enzymes presented in Fig 1 and Fig 3 need to be depicted in 3D in order to evaluate how close they are in 3D. Intermolecular distances should be shown on these figures. To convince the reader that LAF237 sits into and makes contacts with enzymes other than DPP4, we need to see the compound docked into the structure of each enzyme of interest.
I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

**Competing Interests:** No competing interests were disclosed.

---

**Author Response 10 Dec 2014**

**Sandeep Chakraborty**, Tata Institute of Fundamental Research, India

We would like to thank you for taking the time to review this paper, and also for your insightful comments. We also apologize for the inordinate time taken to respond to your comments. A lot of this time was spent in understanding docking methods, instead of blindly applying this to the problem at hand. A by-product of this learning process was the implementation of a new method (DOCLASP) for docking molecules to proteins\(^1\). We have docked vildagliptin to the PI-PLC structure complexed with myo-inositol using DOCLASP. Based on your suggestion, we have also done a comprehensive analysis of all 76 known DPP4 structures liganded to inhibitors till date.

Please find out detailed responses to your comments below.

- **The successful targeting of DPP4 using small molecule compounds to treat type 2 diabetes has attracted a great deal of attention towards the study of this protease.** The authors applied sophisticated techniques that they have developed in order to discover that two DPP4 inhibitors, including one that is in limited clinical use, can to some extent inhibit the activity of a bacterial lipase (PI-PLC). Many lipases and esterases and hydrolases including DPP4 and related enzymes use the alpha/beta hydrolase fold and the authors show how this related protein topology can place the residues in positions that are sufficiently similar to interact with an inhibitor. The major difficulty with this paper is that it attempts to connect these data with possible clinical outcomes. No evidence for such a link is presented. Therefore, the title and much of the conclusions need to be modified so that they reflect the data without speculation.

We have tried to keep away from taking sides on the clinical outcomes, since that is not our forte. Also, we believe our title is innocuous in that context - it just speaks of promiscuous scaffolds. We only highlight that if (and only if) our data of PIPLC inhibition holds true for human lipases, then it might provide some arguing points for those worried about the side effects of these drugs.

For example, we say ‘The reported elevated levels of serum lipase, although contested, could be rationalized by inhibition of lipases reported here’. If you could kindly point out specifically any speculations that is unwarranted, we will modify those.

- **Two inhibitors of DPP4, LAF237 and K-579, were studied.** K-579 is not in clinical use. LAF237 is licensed in Europe and is known to exhibit some inhibition of the DPP4-related proteases DPP8 and DPP9. The extent of inhibition of DPP8 and DPP9 by LAF237 is believed not to have physiological effects in humans.

Since this study does not emphasize on the clinical relevance of the inhibitions (but on the methodology of finding such interactions), and we are not a group specializing in diabetes, we believe the choice of the inhibitors would not alter our reasoning our conclusions.
The IC50 of LAF237 on DPP9 is less than 0.01 mM. The IC50 of LAF237 on bacterial PI-PLC is 0.1 mM, which is close to the lower limit of detection of inhibition of an enzyme.

We agree to this point. However, K-579 was inhibiting even at nanomolar concentrations.

No mammalian homolog of PI-PLC was examined.

We are currently evaluating that possibility.

The literature that the authors cite to suggest that DPP4 inhibition might be detrimental for human health, particularly the pancreas, is data on sitagliptin or exenatide. Exenatide is not a DPP4 inhibitor and sitagliptin is quite different to LAF237, both in protease specificity and in chemical structure.

We were referring to the inhibitor part of the data, but that point needs to be made explicit as you have correctly pointed out. Also, we agree that the possible difference of sitagliptin with LAF237 needs to be stated. We have modified the text to include these criticisms. Once again, we reiterate we intend not to comment on clinical outcomes or debates, but to suggest a rational methodology to act as a guide for tests that look for possible interactions.

The contact points of LAF237 and sitagliptin in the catalytic site of DPP4 differ considerably. The authors present no data on sitagliptin or any other DPP4 inhibitor (other than LAF237) that is in the clinic.

We have included a comprehensive study on the contact points of various inhibitors. Once again, this does not negate any of our conclusions.

The images of overlaid catalytic triads of various enzymes presented in Fig 1 and Fig 3 need to be depicted in 3D in order to evaluate how close they are in 3D.

The 3D images of the superimposition of these enzymes are not pleasing to the eye, since they lack structural homology. However, we have added a PyMol script in case someone wishes to do that (Superimposeproteins.p1m). The script specifies the color coding of the residues.

Intermolecular distances should be shown on the figures.

Once again, we think that the intermolecular distances clutter the figure. The superimposition gives an approximate idea of the congruence. The exact values are specified in Table 1. We have modified the legend of Fig.3 to specify that.

To convince the reader that LAF237 sits into and makes contacts with enzymes other than DPP4, we need to see the compound docked into the structure of each enzyme of interest.

As mentioned previously, we have docked sitagliptin to PI-PLC using DOCLASP. We have provided the Pymol script as supplementary data to help visualize the docking. There is no solved structure where LAF237 inhibits DPP4.

Once again, we are thankful for the comments. We hope that we have addressed your concerns by
the changes that we have made, and that the manuscript will be found suitable in the modified form.

References
1. Chakraborty S. DOCLASP - Docking ligands to target proteins using spatial and electrostatic congruence extracted from a known holoenzyme and applying simple geometrical transformations [v2; ref status: awaiting peer review, http://f1000r.es/4pb]

Competing Interests: No competing interests were disclosed.

Referee Report 06 March 2014

doi:10.5256/f1000research.3236.r3818

Rodney Rouse
Division of Applied Regulatory Science, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD, USA

Disclaimer: I lack the protein chemistry expertise to comment on the assumptions and protein chemistry used in the computational method described in this article.

The title and abstract are appropriate. The overall experimental design is simple but strong and well suited for this project. The methods were generally well described. The conclusions are not overstated and any implications are justified based on the presented data. The article is very well written.

This is a very interesting study that uses a previously defined computational method, CataLytic Active Site Prediction (CLASP), that compares structural and charge similarities of catalytic sites to identify functionally similar proteins. This methodology was used to assess the potential for adverse events based on off target effects of the inhibitors of DPP-IV. Using CLASP, the authors had previously indentified a Bacillus cerus phosphoinositide specific phospholipase-C (PI-PLC) as similar in active catalytic site to the enzyme, DPP-IV. They used laboratory techniques to verify this finding.

In the present study, the authors demonstrated the ability of two separate DPP-IV inhibitors to significantly reduce the activity of this PI-PLC in the lab. Subsequent to this experimentation, the authors returned CLASP to identify catalytic sites in other proteins that might also be inhibited by DPP-IV inhibitors thereby yielding unforeseen inhibition and biological effects. As applied to the case of DPP-IV inhibitors, which are not extremely specific, the authors identify a number of other proteins that could be promiscuously impacted by DPP-IV inhibitors thereby providing mechanisms for unexpected adverse events. Although the significance of DPP-IV inhibitor related adverse events has yet to be determined, the fact that changes have been reported non-clinically and clinically are undeniable. Eventually, the benefit of these molecules may far outweigh their associated risks, but the authors provide a potential path forward for investigation of unexpected events with this class of drug. If contradictory reports persist, this path may require further illumination.

The approach is theoretically similar to using structural similarities to identify off target receptor binding and consequent biological effects, an expanding approach in safety assessment and in identification of
mechanisms for adverse events in the pharmaceutical lifecycle. Similarly, this method could be predictive for off target effects and suggest what those effects might be. However, whether this is a method that can be generally applicable to other molecules is beyond my ability to comment and the scope of this work.

Comments/Suggestions:

1) Were the inhibition experiments done in duplicate, triplicate, etc? Some slight expansion of the protocols would help with attempts to replicate.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

**Competing Interests:** No competing interests were disclosed.

---

**Author Response 10 Mar 2014**

**Adela Rendón-Ramirez**, Unidad de Biofísica CSIC UPV, Spain

Dear Dr Rouse,

We would like to thank you for taking the time and reviewing our paper. Your positive comments encourage us to further our research in this area.

We concur with your statement - “Eventually, the benefit of these molecules may far outweigh their associated risks”. And it is our endeavor to improve the accuracy and generality of our method through different compounds. We would specifically like to highlight another case of antagonist binding identified through CLASP, although in this case most alkaline phosphatases were not affected - Chakraborty *et al.* (2012)

The data for PI-PLC inhibition using DPP4 inhibitors, as shown in Figure 2, are average values of two closely similar experiments. We will revise the manuscript to include this point when we hear from another referee.

Best regards,

Sandeep Chakraborty and Adela Rendón-Ramirez

**Competing Interests:** No competing interests were disclosed.