In silico study to predict potential precursors of human dipeptidyl peptidase-IV inhibitors from hazelnut

Congcong Li, Renrui Qi, Jiawei Yuan, Lu Han, Song Wang, Wannan Li and Weiwei Han

Key Laboratory for Molecular Enzymology and Engineering of Ministry of Education, School of Life Science, Jilin University, Changchun, China; Laboratory of Theoretical and Computational Chemistry, Institute of Theoretical Chemistry, Jilin University, Changchun, China

CONTACT Wannan Li (liwannan@jlu.edu.cn); Weiwei Han (weiweihan@jlu.edu.cn)

ABSTRACT

Chinese hazelnut was chosen to become a probable precursor of biological active peptides via computer simulations in this article. There were a large number of bioactive peptides in Chinese hazelnut sequences according to analytical results from the BIOPEP database. The most prominent of these was the inhibitory peptide for dipeptidyl peptidase-IV (DPP-IV; EC 3.4.14.5), which can be used to treat type 2 diabetes, so the theoretical method to obtain DPP-IV inhibitory peptides by hydrolysis with a single or combination of enzymes was studied. Cytotoxicity analysis performed by ToxinPred showed that all of the DPP-IV inhibitory peptides generated from protein hydrolysis were not cytotoxic. Structural interaction fingerprint analysis revealed that Asp663 and Phe357 may be important residues for ligand binding. In order to further understand the inhibitory mechanism of peptide, VR with lowest half maximum inhibitory concentration (IC50) and IPI (inhibitors have been reported) were selected as ligand of DPP-IV to perform steered molecular dynamics simulations and PMF calculations. The results showed that P1 is the preferred (un)binding tunnel for the inhibitors obtained. Our findings help in the development of new DPP-IV inhibitors which were derived from common food.

1. Introduction

Dipeptidyl peptidase-IV (DPP-IV; EC 3.4.14.5) is a serine protease that preferentially cleaves proline or alanine dipeptides from the N-terminus of polypeptides and proteins (Yaron & Naider, 1993). Inhibition of DPP-IV can help maintain the insulino-tropic effect of incretins over a longer period of time. Thus, the use of DPP-IV inhibitors to treat type 2 diabetes considered a novel treatment approach for the disease. Peptide intake can effectively improve personal immunity. Thus, methods to obtain better peptides are a research hotspot. Theoretical methods have been used to design various bioactive peptides from food-based proteins, such as yak milk casein, soy and lupin protein, milk protein isolate (Nongonierma et al., 2018), salom (Li-Chan et al., 2012), Euphausia superba protein (Ji et al., 2017), rice bran protein (Udenigwe, 2016) and amaranth (Valarde-Salcedo et al., 2013), amongst others. Therefore, research on obtaining DPP-IV inhibitor peptides is an important topic in the development of the food industry.

The use of single or combined proteases to hydrolyze proteins is the most common method used to find bioactive peptides from protein molecules. A number of useful theoretical proteolysis tools, such as the BIOPEP database (Piotr et al., 2008) [http://www.uwm.edu.pl/biochemia/biopep/start_biopep.php] and ExPASy Peptide Cutter (Panu et al., 2012) [http://web.expasy.org/peptide_cutter/], have been used to predict bioactive peptides by combining the proteases of specific cleavage sites with various substrates of known protein sequences. Compared with that in actual experiments, the time required to screen bioactive peptides present in different protein sources can be significantly shortened by in silico studies. Theoretical studies predicting the potential precursors of many food-derived proteins is an interesting endeavor with multiple benefits.

Chinese hazelnut (Corylus heterophylla) is a well-known wild hazelnut species that belongs to the family Betulaceae (Liu et al., 2018). Hazelnuts are considered a rich source of several compounds beneficial to human health (Adamo et al., 2017). In silico analysis and bioinformatics-based methods have been proven to be useful for predicting potential proteins as precursors of DPP-IV inhibitory peptides (Bleakley et al., 2017; Nongonierma et al., 2018). Canola, chicken egg, oat and wheat have been identified as potential sources of DPP-IV inhibitory peptides. However, DPP-IV peptides from hazelnut proteins or protein hydrolysates have not been reported.

In this study, four hazelnut proteins (i.e. PHD finger family protein [Corylus heterophylla, GenBank: AEF79998.1] from a number of cold-related genes of Chinese hazelnut; COR413-PM1 [C. heterophylla, GenBank: ADM64617.1] from Chinese hazelnut; WR11 [C. heterophylla × Corylus avellana, GenBank: AVT56113.1], a homolog of WRINKLED1 from hybrid Chinese hazelnut; and CBF/DREB1 transcription factor 1 [C. avellana × C. heterophylla, GenBank: AMK62112.1], a homolog of hybrid...
Chinese hazelnut) were determined. The four sequences obtained were used to predict potential precursors (inhibitors) of human DPP-IV, as shown in Table 1.

This study used in silico methods to predict the potential use of hazelnut as a precursor of bioactive and DPP-IV inhibitory peptides. Nevertheless, the results predicted in silico are not always obtained in vivo and should be validated in vitro and in vivo before taking conclusions, so the results only provided some clues for the experimental validation.

A total of 50 active peptides were obtained by determining the optimal enzyme combination for the efficient production of DPP-IV inhibitory peptides. Structural interaction fingerprint (SiFt) analysis was then conducted to identify important residues that may bind to the inhibitors. Amongst the 50 peptides detected, VR showed the lowest half maximum inhibitory concentration (IC50). Steered molecular dynamics (SMD) simulations and PMF calculations were employed to explore the unbinding pathway of the inhibitors. Ultimately, our results will help in efforts to design novel DPP-IV inhibitors for the cognitive enhancement of ligands.

2. Materials and methods

2.1. Prediction of bioactive peptides derived from hazelnut

The BIOPEP database http://www.uwm.edu.pl/biochemia/biopep/start_biopep.php can be used to evaluate proteins as a potential source of bioactive peptides (Minkiewicz et al., 2008). According to the equation, the potential biological activity of protein fragments was determined, and the following parameter was calculated: $A = \frac{a}{N}$, where $A$ is the frequency of occurrence of fragments with a given activity, $a$ is the number of active fragments in the protein sequence and $N$ is the number of amino acid residues in the protein.

2.2. Release of DPP-IV inhibitory peptides and toxicity prediction

The BIOPEP database can simulate proteolysis and calculate the frequency of release of fragments with a given activity by a selected enzyme (AE) (Minkiewicz et al., 2011). The selected hazelnut sequences were subjected to hydrolysis in silico using five different enzymes (i.e. chymotrypsin, trypsin, papain, thermolysin and proteinase K). Single enzymes or two- or three-enzyme combinations were applied to obtain the maximum number of inhibitory peptide fragments from hazelnut. ToxinPred (Gupta et al., 2015), which is available at https://webs.iiitd.edu.in/raghava/toxinpred/index.html, was used to predict the toxicity of DPP-IV inhibitory peptides derived from the four hazelnut sequences.

2.3. Molecular docking and structural interaction fingerprint to analyze DPP-IV – peptide binding interactions

Chemical peptides optimised at the B3LYP-6-31G* level by using Gaussian 09 software (Xu et al., 2018). Human DPP-IV was deposited to the Protein Data Bank (PDB Code: 3WQH). To validate the reliability of the docking results, IPI was reasonably docked into the active pocket of DPP-IV by CDOCKER (Wu et al., 2003), Autodock Vina (Trott & Olson, 2010), and LeDock (Zhang & Zhao, 2016) three docking softwares before steered molecular dynamics simulations.

Discovery studio 2.5 was used to dock 50 ligands with the protein. The docking method was CDOCKER. The force field used in this docking method was CHARMM force field. Each ligand docking will produce 10 poses. After docking, the docking of 50 ligands was analysed with the lowest energy by using receptor ligand interactions, then the average SiFt will generate.

2.4. Steered molecular dynamics simulation

SMD (Izrailev et al., 1999) is a simulation method based on conventional techniques to study ligands, donor binding and separation and a number of mechanical and energy properties in the simulation process (Zhu et al., 2017). In this research, SMD simulations were performed to drive the dissociation of two inhibitors (i.e. IPI and VR) from the enzyme. During SMD simulation, a time-dependent external force was applied to the inhibitor in a specific direction so that it is eventually separated from its receptor. The two directions were used in Nanoscale Molecular Dynamics (NAMD) (Kalé et al., 2011) 2.10b1 version software using the CHARMM27 all-atom force field (Klauda et al., 2010). One of the pulling directions was set to the initial location of the inhibitor in the active site and the center of the $\alpha$ atoms of Gln320 and Arg596. The second direction was set to the initial location of the inhibitor in the active site and the center of the $\alpha$ atoms of Tyr547 and Val711. The pulling velocity was set to 0.01 Å ps$^{-1}$, which was slower than some of the velocities used in other nuclear receptors by the same method. SMD simulations were performed from the snapshot structure at 10 ns.

2.5. PMF calculations

In order to quantitatively visualize the potential energy changes of the two substrates, we calculated the change in PMF along the egress pathway mapped out from the SMD simulations. The two substrates featured three different RCs, which referred to the distances between the center of mass of each substrate and the residue oriented toward the opening of the pathway. The ABF method of NAMD software can be used to calculate the relevant energies (Minoukadeh et al., 2010). Our study divided these RCs into numerous small windows, each of which was simulated for 4 ns. Finally, a completion of all these parated simulations belonging to the same RC.

3. Results and discussion

3.1. Frequency of occurrence of bioactive fragments in hazelnut

The BIOPEP database showed the frequency of occurrence of bioactive fragments for each type of bioactivity in the
hazelnut sequences. The amino acid length and ADPP-IV values used in this study are listed in Table 1. As seen in Table 1, the A values of DPP-IV inhibitory peptides determined by in silico analysis are 0.58 (PHD finger family protein), 0.6980 (COR413-PM1), 0.6330 (WR1) and 0.7240 (CBF/DREB1 transcription factor 1). CBF/DREB1 transcription factor 1 has a higher \(\Sigma A\)
value (1.39) than the other peptides (PHD finger family protein, 1.12; COR413-PM1, 1.39; WRI1, 1.30). Four of these classes were found in all analysed hazelnut sequences, such as DPP-IV inhibition, ACE inhibition, immunomodulating, antioxidative and stimulating glucose uptake activity. Analysis of the frequency of occurrence of fragments with a specific activity indicated that hazelnut contains a large number of DPP-IV inhibitors in each of its sequences (Figure 1).

3.2. In silico proteolysis of hazelnut for the production of DPP-IV inhibitory peptides

Analysis of the biological activity of the hazelnut sequences indicated that the nut is a suitable precursor for the synthesis of various bioactive peptides. The cleavage sites of chymotrypsin, trypsin, proteinase K, thermolysin and papain selected in this study are shown in supplementary material Table S1. Five proteases were used to hydrolyze hazelnut for the production of DPP-IV inhibition peptides in silico. Hazelnut was first hydrolysed with a single enzyme, followed by a two-enzyme combination and then a three-enzyme hydrolysed hazelnut, two-enzyme hydrolysed hazelnut yielded higher AE values. The most effective combination for the production of DPP-IV inhibitory peptides determined by in silico proteolysis in all two-enzyme combinations was chymotrypsin + papain, which resulted in a maximum AE value of 0.4264. The DPP-IV inhibitory peptides of hazelnut were obtained by in silico proteolysis using the enzyme combination chymotrypsin + papain. The majority of DPP-IV inhibitory peptides produced by chymotrypsin hydrolysis can be retained by addition of the second enzyme, papain, which can produce new DPP-IV inhibitory peptides without damaging the original peptides produced by chymotrypsin. Therefore, the combination of these two enzymes could produce the most effective DPP-IV inhibitory peptides amongst the treatments analysed. Interest in using combinations of enzymes to hydrolyze different proteins for the production of DPP-IV inhibitory peptides has increased in recent years. In silico simulation of enzyme combinations for hydrolysis can help guide bench experiments to retain the greatest number of the most bioactive peptide fragments. However, proteolysis of protein in silico have its negative aspects, the peptides generated in silico cannot always be generated in vitro, so the predicted results only can provide basis for experiments and save time.

3.3. Toxicity prediction for DPP-IV inhibitory peptides generated from hazelnut

ToxinPred is an in silico method to predict peptide/protein toxicity. Firstly, a prediction model was established based on the small molecular toxic peptides obtained from different databases, meanwhile, SwissProt and TrEMBL databases provided non-toxic peptides. Using machine learning techniques, the developers developed a quantitative matrix model based on the distribution of toxic peptides to predict the toxicity of peptides according to their properties. The accuracy rate of the results were about 90% (Gupta et al., 2013). The support vector machine (SVM) scores of the 50 peptides yielded negative values, thereby indicating their nontoxicity.
Interestingly, the SVM score of VR (Val–Arg) was −0.80. This result reveals that the peptide may be potentially used as a natural active ingredient of antidiabetic functional foods or drugs. Indeed, finding safe antidiabetic compounds from natural food sources has become a strong driving force behind new research in food processing, functional food development and treatment.

### 3.4. Molecular docking and structural interaction fingerprint studies

After molecular docking was completed, the rationality of the results were evaluated by docking score, if the root mean square deviation (RMSD) of the heavy atoms in the ligand (substrate) less than 2 Å, which was regarded as the docking result within a reasonable range, otherwise, the docking result was unreasonable. The redocking results were shown in supplementary material Figure S1, it was suggested that the results of CDOCKER docking are more reasonable, and the RMSD value of the resulting conformation relative to the original crystal structure was 1.26 Å, while for Autodock

---

![Figure 4.](image) (a) The predicted tunnels. (b) The active residues around tunnel 1. (c) The active residues around tunnel 2.

### Table 3. Characteristics of the top ranked tunnels of DPP identified by CAVER 3.0 in molecular dynamics trajectory using the probe radius of 1.5 Å and the clustering threshold of 8.5.

| Rank | Pathway cluster | No_snaps | Average bottleneck radius (Å) | Average tunnel length (Å) | Average tunnel curvature (Å) | Priority |
|------|-----------------|----------|-------------------------------|---------------------------|-------------------------------|----------|
| 1    | P1              | 1000     | 3.18                          | 22.14                     | 1.19                          | 0.91     |
| 2    | P2              | 1000     | 3.15                          | 45.68                     | 1.25                          | 0.83     |
| 3    | P3              | 459      | 2.11                          | 38.82                     | 1.73                          | 0.32     |
| 4    | P4              | 110      | 1.61                          | 41.36                     | 1.32                          | 0.06     |
| 5    | P5              | 31       | 1.57                          | 47.96                     | 1.65                          | 0.02     |
| 6    | P6              | 15       | 1.61                          | 32.89                     | 1.84                          | 0.01     |
| 7    | P7              | 9        | 1.93                          | 55.05                     | 1.39                          | 0.01     |
| 8    | P8              | 6        | 2.43                          | 48.48                     | 1.96                          | 0.001    |
| 9    | P9              | 2        | 1.51                          | 56.15                     | 1.29                          | 0.001    |
| 10   | P10             | 1        | 1.79                          | 62.28                     | 2.20                          | 0.001    |

---

### Table 4. Bottleneck residues of the top ranked tunnels of DPP identified by CAVER 3.0 in molecular dynamics trajectory using the probe radius of 1.5 Å and the clustering threshold of 8.5.

| Pathway | P1 (%) | P2 (%) |
|---------|--------|--------|
| Bottleneck Residues |        |        |
| H740    | 100%   | K463(100%) |
| N710    | 100%   | S462(100%) |
| S630    | 100%   | S209(100%) |
| Y547    | 100%   | E206(100%) |
| E205    | 100%   | E205(100%) |
| R125    | 100%   | P218(100%) |
| E206    | 99.7%  | L410(99.8%) |
| W620    | 99.3%  | H363(99.8%) |
| N51     | 99%    | R125(98.7%) |
| Y666    | 98.9%  | F357(92.5%) |
| Y48(89.6%) |       | W62 (85.3%) |
| D708(55.3%) |       | S158(74.7%) |
Vina and LeDock, the RMSD values was 2.01 and 4.15 Å, respectively.

Several homo sapiens, low-resolution, DPP-IV-ligand complex structures were selected from the database as the alternative structure, then molecular docking was performed to compare the affinity between them. The molecular docking results of the different DPP-IV crystal structures and ligands were listed in supplementary material Table S2. Obviously, 3WQH showed good affinity, so it was chosen as the structure for the further simulations.

To date, the specific interactions between peptides from hazelnut and DPP-IV remain unclear. Peptides can be hydrolysed by DPP-IV and show competitive inhibition. According to a proposed reaction mechanism and related reports (Watanabe et al., 2015), the active site obtained from molecular docking nearly covered all the known important residues, including the catalytic triad Ser630, Asp708 and His740.

In our study, the docking were performed for 50 peptides using CDOCKER (Wu et al., 2003). The docking results yielded the 10 top poses of the compound at the active site, and the conformation with the lowest docking score was selected. A seen in Table 2, all ligands were docked in the active site. The position and volume of the binding site were determined by analysing the binding patterns of different molecules. The molecular surface representations showed the interaction of inhibitors with DPP-IV (Figure 3). Glu205, Glu206, Tyr662, Arg125, Asp663, Ser630, Tyr547, Tyr666, His740, Phe357, Tyr631, Asn710, Arg669, Trp629, Arg358, Ser209 and His126 were important residues for peptide binding. Ser630 and His740 are part of the catalytic triad. Two highly conserved glutamic acid residues in the predicted β-propeller domain of DPP-IV are required for its enzyme activity and form salt bridges with the N-terminus (P2 residue) of a peptide substrate located on a short helix insertion in blade 4 of the β-propeller domain. The side chain of Arg125, which is located in blade 2 of the β-propeller domain, is involved in the recognition of the carbonyl group of a peptide substrate. Tyr547, Tyr662 and Tyr666 from the α/β hydrolase domain may participate in hydrophobic interactions with the P1-binding pocket for smaller side chains (Pro

Figure 5. (a) The typical force profiles of IPI pulling out of the DPP-IV pocket along P1 and P2. (b) The typical force profiles of VR pulling out of the DPP-IV pocket along P1 and P2. (c) The force profiles comparison of IPI and VR pulling out of the DPP-IV pocket along P1.

Figure 6. (a) The docking pose of DPP-IV with IPI. (b) The docking pose of DPP-IV with VR.
or Ala) of the substrate. Our SIFT results are consistent with the experimental data. Approximately 48% of the inhibitors can participate in charge interactions with Asp663, and 21% of the inhibitors can participate in hydrophobic interactions with Phe357. Thus, Asp663 and Phe357 are likely to be important for inhibitor binding.

3.5. Steered molecular dynamics studies

3.5.1. CAVER calculations

All tunnels amongst the top ranked pathway clusters were identified using CAVER 3.0 (Chovancova et al., 2012). DPP-IV could theoretically have two (un)binding tunnels, namely, P1 and P2, for inhibitors (Figure 4(a)). Table 3 indicated that the P1 tunnel was the dominant pathway because it was the most frequently identified collective pathway with the highest mean radii of bottlenecks. Figure 4(b) and Table 4 revealed that the most frequent bottleneck residues in the DPP-IV P1 tunnel were formed by His740 (100%), Asn710 (100%), Ser630 (100%), Tyr547 (100%), Asn51 (100%), Glu205 (100%), Arg125 (100%), Glu206 (99.7%), Trp629 (99.3%), Tyr48 (89.6%) and Asp708 (55.3%). By comparison, the bottleneck residues in the DPP-IV P2 tunnel include Arg463 (100%), Ser462 (100%), Ser209 (100%), Glu206 (100%), Glu205 (100%), Pro218 (100%), Leu410 (99.8%), His363.
(99.7%), Phe351 (92.5%), Trp62 (85.3%) and Ser158 (74.7%). These residues may be concluded to play important roles in the unbinding pathway via the P1 and P2 tunnels.

3.5.2. Inhibitor unbinding pathway via P1
We conducted SMD simulations to estimate the rupture forces of the two most frequently observed unbinding pathways quantitatively. We choose two inhibitors (i.e. VR, the inhibitor with the lowest IC50 in hazelnut, and IPI, the optimal inhibitor for DPP-IV). The VR inhibitors which had the lowest IC50 value were selected from the pool of 50 peptides on the basis of their reported IC50 values. The maximum force ($F_{\text{max}}$) and sum of forces ($F_{\text{sum}}$) extracted from the SMD simulation trajectories were listed in supplementary material Table S3. The SMD simulations for IPI and VR were repeated six times (supplementary material Figure S2). The $F_{\text{max}}$ and $F_{\text{sum}}$ values (IPI and VR) obtained suggested that P1 was the preferred tunnel for inhibitors (Figure 5(a,b)). Figure 5(c) shows that IPI can escape from DPP-IV via P1 with greater ease than VR.

Figure 8. The distance changes between (a) Glu205, (b) Glu206, (c) Asn710, (d) His740 and VR.
P1 was considered the favoured pathway through which inhibitors escape from DPP-IV. Glu205, Glu206, Tyr662, Tyr631, Tyr547 and Arg125 formed hydrogen bonds with IPI (Figure 6(a)). The first peak of IPI via the P1 unbinding

![Figure 9](image_url). The distance changes between (a) His363, (b) Lys463 and IPI

![Figure 10](image_url). The distance changes between (a) Ser158, (b) Lys463 and VR.
pathway occurred at 0.65 ns. Asp663 and Tyr66 formed new hydrogen bonds with IPI (supplementary material Figure S3(a)), and other hydrogen bonds were maintained. The second peak of IPI via the P1 unbinding pathway occurred at 1 ns. The hydrogen bonds formed with Asp663, Tyr666, Tyr662 and Tyr63 disappeared, and new π–alkyl interactions between Phe357 and IPI were formed (supplementary material Figure S3(b)). Glu205, Glu206, Tyr547 and Arg125 also form hydrogen bonds with IPI. The third peak of IPI via the P1 unbinding pathway occurs at 1.61 ns. Arg125, Glu205 and Asn710 form hydrogen bonds with IPI. Trp629 forms a new π–alkyl interaction with IPI (supplementary material Figure S3(c)). In the IPI unbinding pathway with DPP IV via P1, Glu205 and Arg125 form hydrogen bonds with IPI, which may prevent the latter from escaping DPP-IV. The distance changes between Arg125, Glu205, Asn710, Trp629 and IPI were shown in Figure 7. The results of the repeated experiment are shown in the supplementary material Figure S7. Thus Arg125, Glu205, Asn710 and Trp629 are major residues for the P1 unbinding pathway for IPI.

Arg125 and Tyr547 form hydrogen bonds with VR in the binding pocket (Figure 6(b)). The first peak of VR via the P1 unbinding pathway occurs at 0.69 ns. At this time, two hydrogen bonds formed with Arg125 disappeared, and two new hydrogen bonds formed between Ser630, Tyr666 and VR, and formed salt bridge with Glu206 (supplementary material Figure S4(a)). The second peak of VR via the P1 unbinding pathway occurs at 0.87 ns, and the salt bridge between VR and E206 is maintained (supplementary material Figure S4(b)). H710 formed new hydrogen bonds with VR. The third peak of VR via the P1 unbinding pathway occurs at 1.10 ns, which a salt bridge forms between E206, E205 and VR. Asn710 and His740 formed hydrogen bonds with VR (supplementary material Figure S4(c)). These interactions may prevent the release of VR from DPP-IV via P1. The distance changes between Glu206, Glu205, Asn710, His740 and VR were shown in Figure 8. The results of the repeated experiment were shown in the supplementary material Figure S8. Thus, Glu206, Glu205, Asn710 and His740 are the major residues for VR via the P1 unbinding pathway.

3.5.3. Inhibitor unbinding pathway via P2

The first peak of IPI via the P2 unbinding pathway occurs at 1.02 ns (supplementary material Figure S5). Glu205, Glu206, Tyr547 and Arg125 form hydrogen bonds (salt bridge) with IPI. The new hydrogen bonds (salt bridge) between Glu205, Glu206 and IPI may force the first peak of the P2 unbinding pathway of IPI. The second peak of IPI via the P2 unbinding pathway occurs at 1.81 ns (supplementary material Figure S5). The hydrogen bonds between Glu205 and Glu206 and IPI are maintained, but the hydrogen bonds between Tyr547 and IPI are broken. Thereafter, new hydrogen bonds between Ser209 and IPI are formed. The third peak of IPI via the P2 unbinding pathway occurs at 3.74 ns, at which point new hydrogen bonds between His363 and Lys463 and IPI are formed. The distance changes between His363, Lys463 and IPI were shown in Figure 9. The results of the repeated experiment are shown in the supplementary material Figure S9. These interactions may prevent the release of IPI from DPP-IV Thus, His363 and Lys463 are major residues for VR via the P1 unbinding pathway.

Arg125 and Tyr547 form hydrogen bonds with VR in the binding pocket (Figure 6(b)). The first peak of VR via the P1 unbinding pathway occurs at 1.06 ns. At this time, the hydrogen bond between Arg125 and VR was broken, but the hydrogen bond between Tyr547 and VR was maintained. Four hydrogen bonds (and salt bridges) are formed between Glu205, Tyr666, Glu206 and Phe357 with VR (supplementary material Figure S6). The third peak of VR via P2 unbinding pathway occurs at 3.77 ns, during which Ser158 and Lys463

Figure 11. The PMF profiles calculated for IPI unbinding from (a) P1 and (b) P2 pathway, VR unbinding from (c) P1 and (d) P2.
form hydrogen bonds with VR. These interactions (salt bridges) may prevent VR from leaving DPP-IV via P2. The distance changes between Ser158, Lys463 and VR were shown in Figure 10. The results of the repeated experiment are shown in the supplementary material Figure S10. Thus, Ser158 and Lys463 are major residues for VR via the P2 unbinding pathway.

In summary, P1 appeared to be the favoured pathway through which the inhibitors IPI and VR could escape from DPP-IV. Regardless of the pathway, IPI can escape from DPP-IV faster than VR. In the P1 tunnel, Arg125 and Glu205 were major residues for VR and IPI. The possible significance of a novel conserved sequence motif Asp–Trp–(Val/Ile/Leu)–Tyr–Glu–Glu–Glu (DW(V/I/L)YEEE) in the predicted beta propeller domain of DPP-IV-like gene family had been reported, and single amino acid point mutations (E205K and E206L) in this motif were essential for the enzyme activity of human DPP IV.

Two highly conserved glutamic acid residues in the beta propeller domain of DPP-IV were required for its enzyme activity. Our results indicated that Glu205 was the major residue through which inhibitors escape from DPPIV via P1 and that Lys463 was the major residue through which inhibitors escape from DPPIV via P2.

3.6. PMF

The PMF profiles calculated for P2 and P1 were depicted in Figure 11(a–d). The free energy curve revealed interesting insights into the unbinding pathways of IPI and VR. When the inhibitors IPI and VR departed from their initial equilibrium position, the free energy rapidly increases, which could be attributed to the interaction of the inhibitor with the active residue sites. The free energy of unbinding should be calculated to obtain a precise, reliable estimate of the relative preference of ligand unbinding pathways. The approximate free-energy barrier of the inhibitor unbinding process was 32.74 kcal mol\(^{-1}\) for P2 and 41.00 kcal mol\(^{-1}\) for P1. Thus, PMF supported P1 as the preferred inhibitor unbinding pathway.

4. Conclusion

We studied the properties of Chinese hazelnut as a potential precursor of DPP-IV inhibitory peptides using in silico proteolysis and found that the hazelnut sequences are highly homologous to those of soybean. This finding indicated that hazelnut may potentially be used to produce various bioactive peptides. We demonstrated that hazelnuts may be used as a precursor of various bioactive peptides, especially DPP-IV inhibitory peptides, through BIOPEP database analysis. Furthermore, when a DPP-IV inhibitory peptide in silico is prepared by hydrolysing a hazelnut sequence with a combination of proteases, considering the enzyme specificity of each protease was necessary to avoid the subsequent addition of an enzyme that could destroy the originally produced active peptide. The online peptide toxicity prediction tool ToxinPred revealed that DPP-IV inhibitory peptides obtained by in silico proteolysis do not show cytotoxicity. Our SIFt analysis revealed that Asp663 and Phe357 may be important residues for ligand binding. Our results showed that P1 is the preferred (un)binding tunnel for inhibitors. Moreover, Glu205 is the major residue through which inhibitors escape from DPPIV via P1, and Lys463 is the major residue through which inhibitors escape from DPPIV via P2. Our results will be helpful in efforts to design new inhibitors of DPP-IV.

Acknowledgements

This work was conducted at the High-Performance Computing Center of Jilin University.

Disclosure statement

The authors declare no competing financial interests.

Funding

This work was supported by the Overseas Cooperation Project of Jilin Province under Grant [20200801069GH].

References

Adamo, M., Labate, A. M., Ferrulli, A., Macrì, C., Terruzzi, I., & Luzi, L. (2017). Effects of hazelnuts and cocoa on vascular reactivity in healthy subjects: A randomised study. International Journal of Food Sciences & Nutrition, 69(2), 1–12.

Bleakley, S., Hayes, M., Shea, N. O., Gallagher, E., & Lafarga, T. (2017). Predicted release and analysis of novel ACE-I, renin, and DPP-IV inhibitory peptides from common oat (Avena sativa) protein hydrolysates using in silico analysis. Foods, 6(12), 108. https://doi.org/10.3390/foods6120108

Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., Biedermannova, L., Sochor, J., & Damborsky, J. (2012). CAVER 3.0: A tool for the analysis of transport pathways in dynamic protein structures. PLoS Computational Biology, 8(10), e1002708. https://doi.org/10.1371/journal.pcbi.1002708

Durinx, C., Lambeir, A. M., Bosmans, E., Falmagne, J. B., Berghmans, R., Haemers, A., Scharp, S., & Meester, I. D. (2000). Molecular characterization of dipeptidyl peptidase activity in serum: Soluble CD26/dipeptidyl peptidase IV is responsible for the release of X-Pro dipeptides. European Journal of Biochemistry, 267(17), 5608–5613. https://doi.org/10.1046/j.1432-1327.2000.01634.x

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., & Raghava, G. (2015). Peptide toxicity prediction (pp. 143–157). Springer.

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., Raghava, G. P. S., & Discovery, O, S. D (2013). In silico approach for predicting toxicity of peptides and proteins. PLoS One, 8(9), e73957, https://doi.org/10.1371/journal.pone.0073957

Izrailev, S., Stepaniants, S., Izraelit, B., Koszin, D., & Schulten, K. (1999). Steeped molecular dynamics. Springer.

Ji, W., Zhang, C., & Ji, H. (2017). Purification, identification and molecular mechanism of two dipeptidyl peptidase IV (DPP-IV) inhibitory peptides from Antarctic krill (Euphausia superba) protein hydrolysate. Journal of Chromatography B, Analytical Technologies in the Biomedical and Life Sciences, 1064, 56–61. https://doi.org/10.1016/j.jchromb.2017.09.001

Kalé, L., Bhave, A., Bohm, E. J., Phillips, J. C., & Dam, H. V. J. J. (2011). NAnoscale molecular dynamics (NAMD). Encyclopedia of Parallel Computing.
