Food Chemistry 214 (2017) 67–76

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Development of a workflow for screening and identification of α-amylase inhibitory peptides from food source using an integrated Bioinformatics-phage display approach: Case study – Cumin seed

Hwee-Leng Siow [a], Theam Soon Lim [b,*], Chee-Yuen Gan [a,*]

[a] Analytical Biochemistry Research Centre, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia
[b] Institute for Research in Molecular Medicine, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia

ARTICLE INFO

Article history:
Received 5 April 2016
Received in revised form 28 June 2016
Accepted 10 July 2016
Available online 11 July 2016

Keywords:
α-Amylase inhibition
Bioactive peptide discovery
Bioinformatics
Phage display

ABSTRACT

The main objective of this study was to develop an efficient workflow to discover α-amylase inhibitory peptides from cumin seed. A total of 56 unknown peptides was initially found in the cumin seed protein hydrolysate. They were subjected to 2 different *in silico* screenings and 6 peptides were shortlisted. The peptides were then subjected to *in vitro* selection using phage display technique and 3 clones (CSP3, CSP4 and CSP6) showed high affinity in binding α-amylase. These clones were subjected to the inhibitory test and only CSP4 and CSP6 exhibited high inhibitory activity. Therefore, these peptides were chemically synthesized for validation purposes. CSP4 exhibited inhibition of bacterial and human salivary α-amylases with IC_{50} values of 0.11 and 0.04 μmol, respectively, whereas CSP6 was about 0.10 and 0.15 μmol, respectively. Results showed that the strength of each protocol has been successfully combined as deemed fit to enhance the α-amylase inhibitor peptide discovery.

1. Introduction

The classic approach for bioactive peptide discovery is extremely challenging. Owing to a large batch of peptides produced during the hydrolysis process, a series of fractionation and purification steps followed by bioactivity screenings are unavoidable (Udenigwe, 2014). The most tedious part of the process is to identify the peptides, which are exhibiting the bioactivity. A tremendous amount of data generated are required to be analysed. The time and the cost associated with this conventional approach often result in low number of commercial bioactive peptides (Sharma, Singh, & Rana, 2011). In addition, the peptides may not be discovered after an extensive processing, especially when bioactivity is associated with additive or synergistic effects of various components in the hydrolysates (Udenigwe, 2014). Therefore, researchers urgently need a feasible workflow, which would allow them to overcome these challenges by involving minimum in-lab experimental work and cost.

Advances in computational technique have enabled *in silico* method to be projected as a dominant tool to improve the efficiency and the effectiveness of drug discovery research (Zoete, Grosdidier, & Michielin, 2009). A number of Bioinformatic tools have been developed, which are fast, automated and featurely low cost. For examples, the molecular docking system is used in investigation on target-ligand interactions and existing database with algorithm is used to predict how likely the peptide to be bioactive (Mooney, Haslam, Pollastri, & Shields, 2012; Trabuco, Lise, Petsalaki, & Russell, 2012). However, an experimental assessment is required to validate the specific bioactivity potential of the peptides.

Phage display is a combination of molecular techniques to identify the peptides, which interact with a specific target, from a peptide library. Based on a genetic engineering of bacteriophages as well as repeated rounds of antigen-guided selection and phage propagation, this approach offers an *in vitro* selection of the target with any specificities and affinities. These characteristics make the phage display technique a powerful and cost-effective method for identifying peptides which are able to bind to the target with high affinity and specificity (Souriau, Hua, Lefranc, & Weill, 1998). However, this approach can be complicated, demanding and time-consuming, especially the screening process with random peptides or domain libraries. Such weaknesses could be addressed by narrowing down a large number of possible candidates to a practicable number which can be synthesized, purchased and tested using the aforementioned Bioinformatics tools. Therefore, by combining the strength of both Bioinformatics and phage

* Corresponding authors.
E-mail addresses: theamsoon@usm.my (T.S. Lim), cygan@usm.my (C.-Y. Gan).

http://dx.doi.org/10.1016/j.foodchem.2016.07.069
0308-8146/

display approaches, it should be able to enhance the bioactive peptide discovery workflow.

Based on our previous studies (Siow & Gan, 2014, 2016a, 2016b), cumin seed protein is a valuable precursor of bioactive peptides. Such peptide with inhibitory potential against α-amylase can be a promising target in many areas of disease control and treatment, especially for the diabetes, obesity, hyperlipoproteinaemia and hyperlipidaemia, by controlling the kinetics of carbohydrate digestion and monosaccharide absorption (Alagesan, Raghupathi, & Sankarnarayanan, 2012). Therefore, the main objective of this study was to develop an efficient workflow for identifying anti-amylase peptides from cumin (*Cuminum cyminum*) seed using an integrated Bioinformatics-phage display approach. We believed that this fundamental workflow has the capability to screen a large batch of peptides within a short period of time. It is also worth noting that there will be a minimum of in-lab experiments involved in order to discover a novel bioactive peptide. As a direct consequence, the cost involved will also be reduced tremendously.

2. Materials and methods

2.1. Materials

Cumin seed was purchased from local markets in Penang. Protamex, with an activity of 1.5 AU-NH/g solid, was purchased from Novozyme A/S (Bagsvaerd, Denmark). *Bacillus* sp. α-amylase (BA) and human salivary α-amylases (HSA) were purchased from Sigma-Aldrich (St. Louis, MO, USA). *Escherichia* coli (*E. coli*) DH 10β and TG1 was purchased Stratagene (La Jolla, CA, USA). SiMAG-Amine beads were purchased from Chemicell (Berlin, Germany). All other chemicals and reagents used in the experiment were of analytical grade and purchased from Sigma-Aldrich or otherwise mentioned.

2.2. Peptide sequence identification using liquid chromatography–mass spectrometry (LCMS)

The preparation of peptide was performed according to our previous studies. Prior to hydrolysis, cumin seed protein isolate (CSPI) was prepared under the optimum conditions: incubation time of 0.6 h; incubation temperature of 26.3 °C; and substrate-to-enzyme (S/E) ratio of 20 w/w. (Siow & Gan, 2014). The peptide was subsequently obtained using enzyme Protamex with the optimized hydrolysis parameters: incubation temperature of 42.6 °C; incubation time of 1.83 h; and S/E ratio of 20 w/w (Siow & Gan, 2016b). The resulting peptides were fractioned and further subjected to LC/MS and MS/MS analyses using Thermo LTQ-OrbitrapVelos (Thermo Scientific, San Jose, CA, USA) coupled with an Easy-nLC II system (Thermo Scientific) to identify the peptide sequence (Siow & Gan, 2016a). Data was acquired using Xcalibur version 2.1 (Thermo Scientific) with a mass tolerance threshold of 5 ppm. De novo sequencing was performed on all datasets using PEAKS studio version 6.0 (Bioinformatics Solutions, Waterloo, ON, Canada).

2.3. Proposed workflow for anti-amylase peptide discovery at identification phase

Generally, the proposed workflow of discovery of α-amylase inhibitory peptides at identification phase, was divided into four main phases: phase 1, *in silico* screening (Section 2.3.1); phase 2, *in vitro* selection (Section 2.3.2); phase 3, *in vitro* validation (2.3.3) and phase 4, peptide-binding site search (2.3.4). The methods used in each phase of workflow were further described in the following sections.

2.3.1. Phase 1: In silico screening

2.3.1.1. Bioactive databases. The mass spectrometry data were searched through the databases like BIOPEP, PeptideDB for the bioactive peptide sequence identification.

2.3.1.2. PeptideRanker. The resulting peptides were submitted to the PeptideRanker web server to predict the probability of bioactivity for each peptide (Mooney et al., 2012), which can be accessed at http://bioware.ucd.ie/. The potential bioactive peptides were selected at a threshold of 0.8.

2.3.1.3. PepSite 2. Protein-peptide interaction was predicted using PepSite 2 web server (Trabuco et al., 2012), which can be accessed at http://pepsite2.russelllab.org. *Bacillus subtilis* α-amylase from the RCSB Protein Data Bank (ID number: 1BAG) was used as enzyme model. The model reliability was determined based on a statistical significant ($p < 0.25$). For peptide with more than 10 amino acid residues in length, it was divided into several segments for the analysis, and its *p*-value was determined based on the average score of each segment.

2.3.2. Phase 2: In vitro selection: Phage display

2.3.2.1. Molecular cloning and DNA sequencing. Amino acids were first deduced from nucleotide sequences by EMBOSS Backtranseq websites (www.ebi.ac.uk). The peptide sequences were codon optimized to obtain a high level of expression in *E. coli* using online codon optimization tools (Integrated DNA Technologies; www.idtdna.com). The gene encoding the peptide sequences was constructed using the synthetic oligonucleotides that were synthesized by First BASE Laboratories Sdn. Bhd. (Seri Kembangan, Selangor, Malaysia). The oligonucleotides were annealed and extended by Klenow fragment (New England Biolabs, Ipswich, MA, US). The synthetic oligonucleotides were cloned into pLABEL phagemid vector derived from pFAB1 (Loh, Leong, Tye, Choong, & Lim, 2015). The extended annealed duplex and phagemid vector were excised with the same restriction enzymes, *Nco1* and *Not1* at the recommended digest condition as provided by the manufacturer (New England Biolabs). The digested annealed duplex and vector were then purified and ligated. The resulting construct was subsequently transformed in the *E. coli* DH10β. Recombinant clones were grown overnight at 37 °C in 2YT agar plate containing 2% glucose and 100 μg/mL ampicillin. Colony polymerase chain reaction (PCR) was employed to confirm the presence of target gene under the following PCR program: 95 °C for 90 s, then with a 20 cycles at 95 °C for 30 s, 55 °C for 30 s, 72 °C for 45 s, and final extension at 72 °C for 300 s. Aliquots of clones that showing the expected size of insert DNA fragments in 1.5% agarose gel were purified using a QIAprep spin miniprep kit (Qiaqen Inc., Chatsworth, CA, USA). The purified products were then sequenced with an LMB3 forward primer (5′-CAGGAAACAGCTATGAC-3′) using the Dye Terminator sequencing method with a high-throughput ABI 3730xl Genetic Analyzer (Centre for Chemical Biology, Universiti Sains Malaysia, Penang, Malaysia).

2.3.2.2. Phage packaging and ELISA. The purified phagemids containing the inserted peptide sequences were then transformed into TG1 cells for phage packaging. Coinfection with M13K07 helper phage (New England Biolabs) (10^{10} colony-forming unit (cfu)/mL) was induced by adding into the exponentially growing TG1 culture (OD_{600} of 0.5) and was further grown for 30 min at 37 °C without shaking. Packaging of phage was carried out by culturing the cells overnight at 30 °C with constant shaking at 180 rpm. Following, a phage based enzyme-linked immunosorbant assay (ELISA) for