Food Chemistry 232 (2017) 673–682

Contents lists available at ScienceDirect

Food Chemistry

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

Milk bioactive peptide database: A comprehensive database of milk protein-derived bioactive peptides and novel visualization

Søren Drud Nielsen, Robert L. Beverly, Yunyao Qu, David C. Dallas *

Nutrition Program, School of Biological and Population Health Sciences, College of Public Health and Human Sciences, Oregon State University, United States

ARTICLE INFO

Article history:
Received 16 December 2016
Received in revised form 4 April 2017
Accepted 9 April 2017
Available online 12 April 2017

Keywords:
ACE
Peptidomics
Digestome
DPP-IV
Casein
Whey

ABSTRACT

During processing and digestion, milk proteins are disassembled into peptides with an array of biological functions, including antimicrobial, angiotensin-converting enzyme inhibition, antioxidant, opioid, and immunomodulation. These functions are summarized in numerous reviews, yet information on which peptides have which functions remains scattered across hundreds of research articles. We systematically searched the literature for all instances of bioactive peptides derived from milk proteins from any mammalian source. The data were compiled into a comprehensive database, which can be used to search for specific functions, peptides, or proteins (http://mbpdb.nws.oregonstate.edu). To review this large dataset, the bioactive peptides reported in the literature were visually mapped on the parent protein sequences, providing information on sites with highest abundance of bioactive peptides.

1. Introduction

Milk serves as the primary nutritional source for the mammalian neonate. Beyond basic nutrients, milk provides an array of functional compounds, including bioactive proteins. Degradation of milk proteins releases peptide fragments that have an array of biological effects that can be different from those of the parent protein. These functional peptides are derived from both casein (including α_{s1}-, α_{s2}-, β-, and κ-casein) and whey proteins (including α-lactalbumin, β-lactoglobulin, and lactoferrin). These peptides have an array of activity, including antimicrobial, angiotensin-converting enzyme (ACE) inhibition, dipeptidyl peptidase IV (DPP-IV) inhibition, opioid agonist and antagonist activities, immunomodulation, mineral binding, and antioxidative functions.

Peptides in breast milk or dairy products can be released from milk proteins by native proteases (Dallas, Murray, & Gan, 2015), during production techniques such as fermentation, and during digestion. Whether these peptides exert bioactive effects depends on whether they reach their site of action. In the gut, for example, opioid peptides can bind to gut opioid receptors and alter gastrointestinal motility, and antimicrobial peptides can inhibit pathogen growth. Peptides that are absorbed into the bloodstream can act systemically; for example, the ACE-inhibitory peptides can lower blood pressure. Indeed, some milk peptides are absorbed into the bloodstream as fragments of κ- and α_{s1}-casein were detected in the blood of adult humans for up to eight hours after milk or yogurt consumption (Chabance et al., 1998).

Functional milk peptides may be uniquely suited for applications as added food ingredients, supplements, or drugs, as they likely have few negative side effects due to the evolution of milk as food for the mammalian neonate. Peptides in general are increasingly being investigated for their therapeutic potential, as they are often safer and more selective than small-molecule drugs, which can have wide-ranging side effects. Indeed, over 60 peptides are now approved by the Food and Drug Administration and available on the market as therapeutics (Fosgerau & Hoffmann, 2015). Peptides as antimicrobials are particularly attractive as therapeutics as their mechanisms of action resist development of microbial resistance (Jenssen, Hamill, & Hancock, 2006). However, the application of functional milk peptides as therapeutics remains a rarity.

Advances in mass spectrometry (increased sensitivity, dynamic range, and spectral acquisition rate) allow for increasingly comprehensive data collection such that thousands of unique peptides can now be identified from biological sources. These technological advances create an urgent need for a comprehensive bioactive peptide database with which to compare the thousands of identified peptides. Each new peptide discovered must be compared with the entirety of known functional peptides in order to determine which have potential *in vivo* activity. Bioactive milk peptides have

* Corresponding author at: Nutrition Program, School of Biological and Population Health Sciences, College of Public Health and Human Sciences, Oregon State University, 118G Milam Hall, Corvallis, OR 97331, United States.
E-mail address: dave.dallas@oregonstate.edu (D.C. Dallas).

http://dx.doi.org/10.1016/j.foodchem.2017.04.056
0308-8146/

been reviewed numerous times over recent decades, but no comprehensive database of these peptides exists. A partial database (milk AMP) does exist (Théolier, Jérémie, Ismail, Julie, & Riadh, 2013); however, it includes only antimicrobial peptides and was last updated September 2012. Some milk peptides are present in generic bioactive peptide databases such as the Antihypertensive Peptide Database (AHTPDB) (Kumar et al., 2015) and BIOPEP (Minkiewicz, Dziuba, Iwaniak, Dziuba, & Darewicz, 2008); however, these databases were far from comprehensive for milk protein-derived peptides and have not been updated within recent years to our knowledge.

The primary aim of the present study was to compile a comprehensive database of functional peptides in milk from mammalian species across the available literature sources. As more and more peptides are discovered, the process of comparing new peptides to previously discovered becomes essential in determining their possible bioactivity. In order to assist with the identification and analysis of novel bioactive peptides, we have constructed a comprehensive functional milk peptide database that allows for examinations of patterns in the data of bioactive peptides: a task that is currently difficult as the data are scattered across hundreds of articles. Our database offers improvement over others in existence in three key areas: it is specific and comprehensive to all milk bioactive peptides across species and proteins, every peptide entry has been thoroughly evaluated in the literature before its inclusion, and it contains several advanced search functions to assist researchers in comparing and analyzing peptidomic data. Using our database, we explored how knowledge of bioactive peptides from the milk of one species can be used to predict yet unidentified functional peptides in the milk of other species. We also formulated visualizations to demonstrate which sites within the milk protein sequences have the most numerous or most potent bioactive peptides to help guide future bioactive peptide research. This database will enable comparison of known functional peptides with biological datasets to explore which bioactive peptides are present in food sources and at various digestive sites.

2. Materials and methods

2.1. Literature search

We constructed our database after searching Web of Science (www.webofknowledge.com) for research articles identifying milk-protein-derived peptides with a biological function. The search was conducted with the terms "biological function" AND peptide AND milk, casein, or whey. The "biological function" terms used were "bioactive" (returned 1099 articles), "antimicrobial" (570 articles), "antihypertensive" (682 articles), "immunomodulatory" (98 articles), "anti-inflammatory" (79 articles), "opioid" (224 articles), "dipeptidyl peptidase IV" (49 articles), "anticancer" (45 articles), and "hypocholesterolemic" (43 articles). We did not specify species in the search terms as we wanted to ensure we identified all bioactive milk peptides from every species that has been investigated.

The search was performed between May and December 2016. We refined this search to only include primary research articles. Each abstract was read, and those describing the biological activity of specific milk peptides were saved in a marked list on Web of Science. Based on the abstracts, we identified original research articles to further review for the identification of peptides with biological function. In total, we identified 258 original research articles that described the biological activity of specific peptides. The milkAMP database (Théolier et al., 2013), AHTPDB (Kumar et al., 2015), and BIOPEP database (Minkiewicz et al., 2008) were also searched for any peptides not found using the literature search. For each entry from these three databases, the identified reference was read and the associated data compiled so that our final database cites only original research articles. Those database entries without appropriate references, describing hydrolysates (a mixture of peptides), or with a modified amino acid sequence were excluded. ACE-inhibitory and DPP-IV inhibitory peptides with a half-maximal inhibitory concentration (IC_{50}) value above 1000 μM were excluded from our database as such peptides represent weak activity and a specific IC_{50} value was often not provided for these instances in the articles (Hernandez-Ledesma, Quiros, Amigo, & Recio, 2007). Several studies identified the bioactivity of milk protein hydrolysates. Of these hydrolysate studies, only those that identified the exact peptide sequences responsible for the bioactive effect were included in the database.

Peptides were mapped to the parent sequence of the major proteins in human and cow milk proteins using an in-house tool (PepEx) (Guerrero et al., 2014). The sequences of specific proteins from different species were aligned using the Protein BLAST alignment tool at blast.ncbi.nlm.nih.gov. Peptides in the database were searched for homology within the same protein of origin but across human, cow, goat, and sheep species with an in-house tool (https://github.com/sdrudn/Peptide_Homology) that returns those peptides with ≥80% identical amino acids (identity score). A second function in the tool was also able to count similar amino acids (D = N, I = L = V and Q = E) as identical.

2.2. Online database

The database is SQLite 3.7.17 on a CentOS 7.1.1503 server. The front-end of the site was developed using HTML, Python 2.7.5 and Django 1.9.7 and is served by Apache 2.4.6. The back-end scripts were written in PERL and Python and use Blast+ 2.5.0.

3. Results and discussion

3.1. Online database

The information retrieved from exploring the literature was used to build an online database of human milk and dairy-derived bioactive peptides (Milk Bioactive Peptide Database, MBPDB, http://mbpdb.nws.oregonstate.edu/). MBPDB improves on other databases in three areas. Unlike previous databases, MBPDB is comprehensive for all bioactive milk peptides regardless of species, protein, or function. As milk is likely the most studied source of bioactive peptides with hundreds to thousands of articles, a database specific to it is a necessary tool for identification of these peptides in peptidomics datasets and comparison with newly identified bioactive peptides. Other databases are restricted to a specific peptide function, such as the milkAMP database which only includes antimicrobial peptides. Others still cover bioactive peptides from a wide variety of plant and animal sources and with different functions, and are difficult to search through for peptides of a specific source or function.

MBPDB also has inclusion limits for bioactive peptides that other databases lack. Every new entry is reviewed before inclusion to ensure it meets the proper standards. Adding a new database entry for a bioactive peptide requires the following information: the unique protein ID (e.g. P02666) from which the peptide is derived, the peptide sequence, the function, the specific activity of that function (optional), and the original research article reference given as title, author and digital object identifier (DOI). Based on the input information, the database automatically adds the numeric start and end position of the peptide in the protein sequence (including the signal peptide) and the species of origin information. The database identifies errors between protein ID