

ConjuPepDB: a database of peptide–drug conjugates

Balázs Balogh^{1,†}, Márton Ivánczi^{1,†}, Bilal Nizami^{2,*†}, Tamás Beke-Somfai² and István M. Mándity^{1,3,*}

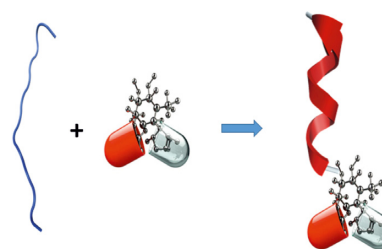
¹Institute of Organic Chemistry, Semmelweis University, H-1092 Budapest, Hógyes Endre u. 7, Hungary, ²Biomolecular Self-Assembly Research Group, Institute of Materials and Environmental Chemistry, Research Centre for Natural Sciences, H-1117 Budapest, Magyar Tudósok krt. 2, Hungary and ³TTK Lendület Artificial Transporter Research Group, Institute of Materials and Environmental Chemistry, Research Centre for Natural Sciences, H-1117 Budapest, Magyar Tudósok krt. 2, Hungary

Received August 13, 2020; Revised September 29, 2020; Editorial Decision October 06, 2020; Accepted October 08, 2020

ABSTRACT

Peptide–drug conjugates are organic molecules composed of (i) a small drug molecule, (ii) a peptide and (iii) a linker. The drug molecule is mandatory for the biological action, however, its efficacy can be enhanced by targeted delivery, which often also reduces unwanted side effects. For site-specificity the peptide part is mainly responsible. The linker attaches chemically the drug to the peptide, but it could also be biodegradable which ensures controlled liberation of the small drug. Despite the importance of the field, there is no public comprehensive database on these species. Herein we describe ConjuPepDB, a freely available, fully annotated and manually curated database of peptide drug conjugates. ConjuPepDB contains basic information about the entries, e.g. CAS number. Furthermore, it also implies their biomedical application and the type of chemical conjugation employed. It covers more than 1600 conjugates from ~230 publications. The web-interface is user-friendly, intuitive, and useable on several devices, e.g. phones, tablets, PCs. The webpage allows the user to search for content using numerous criteria, chemical structure and a help page is also provided. Besides giving quick insight for newcomers, ConjuPepDB is hoped to be also helpful for researchers from various related fields. The database is accessible at: <https://conjupepdb.ttk.hu/>.

GRAPHICAL ABSTRACT



INTRODUCTION

The covalent binding of a small organic drug molecule to a peptide provides the peptide–drug conjugate (1–3). In general, the small organic drug molecule in itself has a prominent, strong pharmacological action. However, for these compounds the reason why conjugation as strategy is useful is because these compounds often suffer from some distinct drawbacks, e.g. high toxicity, low tissue penetration, etc. Moreover, their site-targeted or site-specific application might also offer substantial benefits.

The peptides used in such conjugates basically do not possess pharmacological action as strong as the small organic drug molecule. They can, however, penetrate membranes easily, and can bind to several receptors. This attribute enables the site-specific targeting of the drug molecule with the concomitant lowering of toxicity.

Peptide–drug conjugates are mainly applied in the field of anticancer therapy (1). In principle, the peptide part is used for the targeted delivery of the anticancer drug molecule, with a lowered systemic toxicity. In several cases, cell penetrating peptides are employed, which can facilitate the appropriate translocation of the anticancer agent across the cell membrane to its desired place of application. Selected examples for drug conjugates in this field are as follows: doxorubicin (Dox), daunorubicin (Dau), methotrexate (Mtx) (4,5) and paclitaxel (taxol) (6) too

*To whom correspondence should be addressed. Tel: +36 1 3826 616; Email: mandity.istvan@pharma.semmelweis-univ.hu
Correspondence may also be addressed to Bilal Nizami. Email: nizamibilal1064@gmail.com

†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

(Figure 1A). Besides improved pharmacological properties, peptide–drug conjugates can also be used in a similar way in diagnostic applications, e.g. for cancer imaging. In this case tumor-selective peptides are conjugated to a fluorescent dye (7). The peptide is prone to bind to a specific receptor of the tumor cell, while the fluorescent dye makes the imaging of the cancer tissue feasible.

A further important field of application is antimicrobial utilisation. Some conjugates were tested on various multiresistant bacteria (8) (Figure 1B), and others were designed to act on viruses (9) (Figure 1C). For this purpose heterocycles and macrocyclic systems were used. Examples with antifungal applications are also known (10) (Figure 1D). Additionally, some conjugates were designed to treat malaria utilising quinoline derivative as antimalarial agent (11) (Figure 1E).

Besides the above areas, increasingly diverse fields of application have also emerged. Anti-inflammatory peptide–drug conjugates were created too. Importantly, by the application of polyanionic peptides, anti-inflammatory drugs were effectively delivered into bones or other types of tumor cells (12) (Figure 1F). Furthermore, osteoarthritis can be treated by some conjugates including methotrexate (13). Some promising applications of peptide–drug conjugates can also be found acting on the central nervous system (14), in fat accumulation inhibition, or in gene delivery (15).

There are some anticancer conjugates which were investigated in clinical trials. One of the most promising compounds is Zoptarelin, which reached phase III trials for metastatic endometrial cancer (16,17). A polyglutamic acid-paclitaxel conjugate was investigated in phase III trials too for the treatment of non-small cell lung cancer (18). Furthermore, Vitafolid, a folic acid-vinca alkaloid conjugated with a peptide moiety was also investigated in phase III trials for the treatment of epithelial ovarian cancer (19).

Careful design principles should be followed for the construction of peptide–drug conjugates (1). First, the drug molecule should be carefully selected. It should have a functional group with which it can be attached to the peptide or to the linker without the loss of bioactivity. Additionally, the linker has to be selected cautiously, thus the length, stability and/or release mechanism, solubility, etc. has to be considered. There are stimuli-responsive and biodegradable linkers too, of which release the active drug molecule into the environment of e.g. the cancerous tissue with which the general toxicity of an anticancer agent can be lowered (1).

The chemistry to bind the small organic drug molecule to the peptides is a challenging task. The chemical bond formed between the two molecules should be orthogonal to other functional groups of the conjugate ensuring the formation of only a single compound. To achieve this, several techniques were applied. In most cases an amide bond (Figure 2A) is formed between the two parts of the conjugate (20). This is basically the simplest chemical approach and can be carried out as a part of the solid phase peptide synthesis (SPPS). Nonetheless, further chemical linkages can be used too, such as sulfonamide bonds derived from sulfonic acids (21) (Figure 2B), triazoles obtained by click reactions (22) (Figure 2C), or carbamates formed using of carbonic acid derivatives (23) (Figure 2D). The utilisation of

ether or ester (Figure 2E, F) and sulfide or disulfide bonds is also possible and feasible (24,25).

Importantly by the conjugation the physicochemical properties of the peptide–drug conjugates can be altered too. By careful design, conjugates with self-assembling property can be gained too and supramolecular hydrogels and further nanostructures can be synthesized. Through this self-association the biomedical properties of the drugs can be fine-tuned too (26,27).

As can be seen, the number of successful examples is rapidly increasing in this field. Considering the important and interesting biomedical and pharmaceutical applications of peptide–drug conjugates, and the relative ease of the chemical synthesis thereof, the construction of a comprehensive database is of considerable current interest. However, according to our knowledge, there is no publicly available database concerning peptide–drug conjugates. It should be noted, that there is a database available for cell penetrating peptides (CPPsite 2.0 Database of cell penetrating peptides <https://webs.iitd.edu.in/raghava/cppsite>) which contains some information on peptide–drug conjugates. Nonetheless, CPPsite focuses on data collected only for cell penetrating peptides (28).

Herein we present ConjuPepDB which comprises data of any conjugates where a peptide and a small organic drug molecule are linked by a covalent chemical bond. It includes information regarding the CAS registration number (CAS RN) of the conjugate, the name of both the peptide and the small organic drug molecule, the biomedical application, and the type of the conjugation. Importantly, a chemical structure search of the small organic drug molecule is implemented too, of which helps significantly to navigate among conjugates. ConjuPepDB is hoped to further increase the progress of this intensively developing field of drug research by providing a quick overview for interested non-specialists. More importantly, detailed pieces of information collected are of particular interest for researchers from various related fields, e.g. chemists, pharmacists, clinicians or biologists.

MATERIALS AND METHODS

Data collection and processing

Each entry in the ConjuPepDB is a manually curated and annotated covalent peptide–drug conjugate collected from the literature. Different keywords such as ‘peptide conjugate’, ‘peptide conjugated’ and ‘peptide conjugations’ combined with ‘drug’ were used to perform the search in the Chemical Abstracts (CAS) database (<https://www.cas.org/>) via the SciFinderⁿ tool (<https://scifinder-n.cas.org/>). A conjugate was included into our database only if it contained a peptide and a small (drug-like) molecule connected through a covalent bond. The latter component, in most cases, was a drug compound (a small-molecule active ingredient) or an imaging compound, but other pharmacologically and biologically relevant conjugates were also found. Each literature entry was manually evaluated and only those with one or more relevant substances (molecules) assigned were included. Peptide conjugates and peptides with protecting groups [such as 9-fluorenylmethyloxycarbonyl

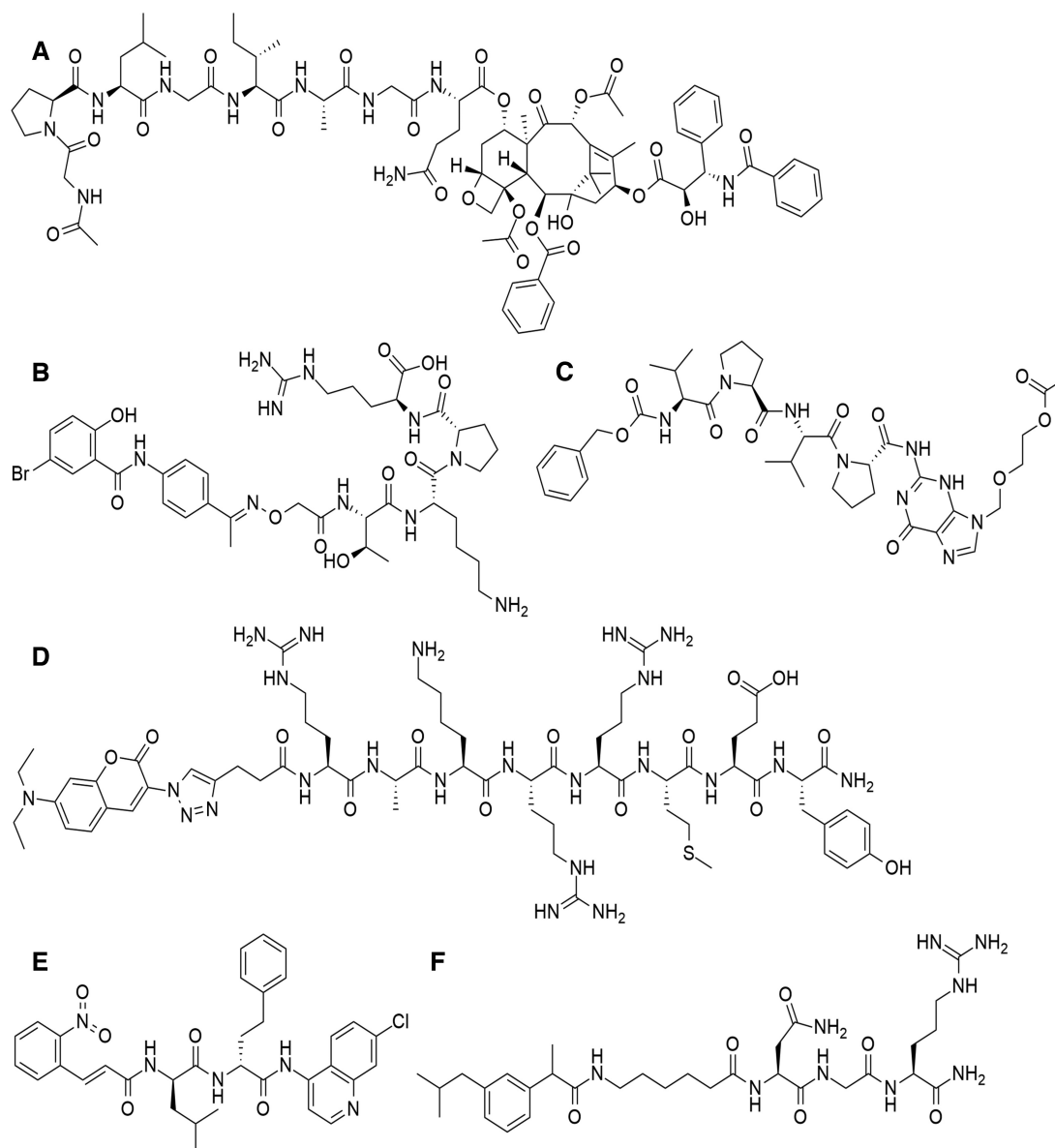


Figure 1. Some selected small organic drug molecules-peptide conjugates examples. (A) paclitaxel (antitumor) conjugate (CAS RN: 1293910-28-7) (6), (B) salicylanilide-derivative (antimycobacterial) conjugate (CAS RN 2100270-47-9) (8), (C) acyclovir (antiviral, HSV-1 and 2) conjugate (CAS RN 1520867-23-5) (9), (D) coumarin derivative (antifungal) conjugate (CAS RN: 1644441-61-1) (10), (E) cinnamic acid/4-aminoquinoline (antimalarial, dual action) conjugate (CAS RN 1392506-21-6) (11), (F) ibuprofen (cancer targeting NSAID) conjugate (CAS RN 2375589-19-6) (12).

(Fmoc)] were not included. A number of articles were omitted, because there were no valid peptide conjugate structures assigned in the CAS database (small molecules only or peptide only or no structures at all) despite indication in their title or abstract. Other sources such as Elsevier's Reaxys database (<https://www.reaxys.com/>) were also searched. However, the number of articles with assigned peptide conjugates were significantly lower (this is the reason why no cross-references were included in our database).

Conjugate structures are referred to by their CAS registration numbers (CAS RN) as well as the small molecule parts. The amino acid sequences of the peptide part are also included. These were taken from the CAS database (if they were specified) or added manually recognised from the 2D

structures. The conjugated small molecules were identified either by name (as referred to within the articles) or manually by their 2D structure. The conjugate structures and the corresponding small molecules were exported from the CAS database in MDL mol format.

According to the type of the covalent bond, each selected conjugate was classified into one of the following groups: amide, urea, triazole, sulfonamide, sulfide, guanidine, ether, ester, disulfide, carbamate, carbon-to-carbon (C-C), azide, ammonium, amidine. Based on the indication of the drug conjugates, they were also grouped into types of biomedical application: anticancer, anti-inflammatory, neurological, antimicrobial while non-drug conjugates were categorised as 'other'. These non-drug conjugates find use-

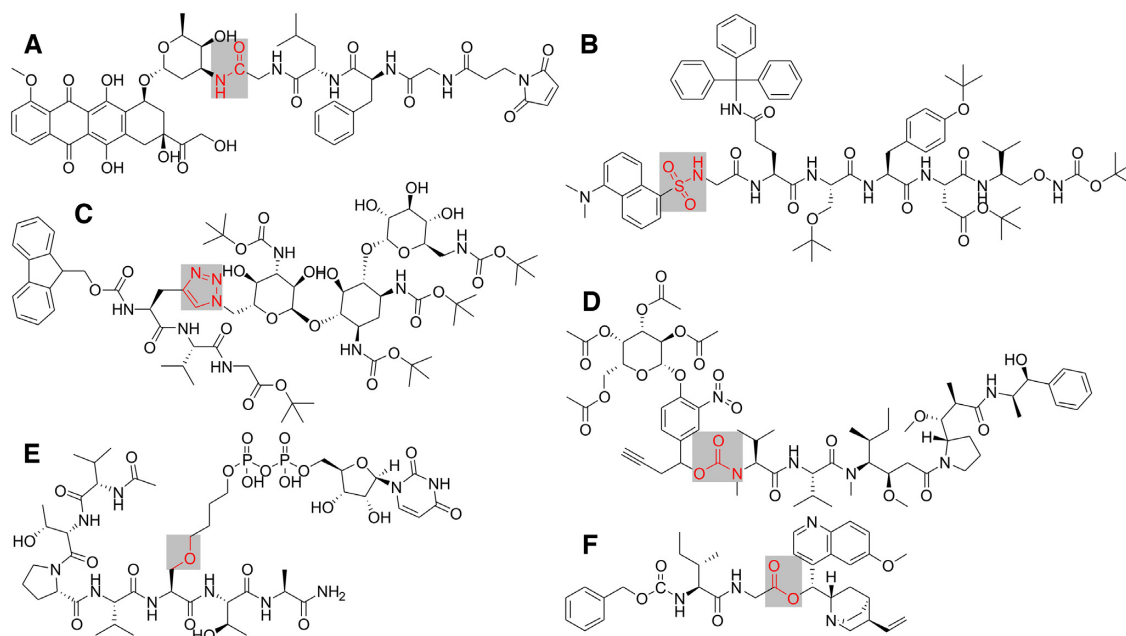


Figure 2. Examples for different chemical linkages. (A) amide (CAS RN 1427500-05-7) (20), (B) sulfonamide (CAS RN 1839533-26-4) (21), (C) triazole, click-reaction (CAS RN 1134204-79-7) (22), (D) carbamate (CAS RN 1418735-54-2) (23), (E) ether (CAS RN 1807977-99-6) (24), (F) ester (CAS RN 1481720-50-6) (25).

ful applications concerning organic synthetic strategy and planning for organic chemists.

Database design and implementation

Information about the peptide–drug conjugates are stored in several relational tables connected to each other via parent-child relationship entity. Apache HTTPS server 2.4 with MySQL server 5.7 in the back-end was used to host the ConjuPepDB database. A web graphical interface is provided for users to view and interact with that data. This dynamic interface is designed using PHP 7.2, HTML5, CSS and JavaScript technologies. Additionally, Bootstrap3 and jQuery libraries are utilised to make a responsive and mobile-first front-end (Figure 3).

JpGraph library (<https://jgraph.net/>) was used to plot charts. Jmol (<http://www.jmol.org/>) is used for rendering the 3D model of the peptide–drug conjugates whereas JSME Molecular Editor v2017-02-26 (29) was employed for depicting the 2D model of small molecules present in a conjugate. Chemical search is implemented using RDKit: an Open-source toolkit for cheminformatics (<http://www.rdkit.org>). ConjuPepDB is built upon relational database management system (RDBMS) technology for easy retrieval and scalability.

Structure search

ConjuPepDB gives users the option to perform different kinds of structure search on the small molecule of the peptide–drug conjugates. A similar method as described in our earlier database (30) has been implemented here. Users can draw the query structures using JSME, a free molecule editor written in JavaScript (29). RDKit toolkit was used

to generate the chemical fingerprints as well as to perform different types of chemical search in the database.

The types of chemical search implemented are as follows:

1. **Exact match.** Returns whether or not two molecules are the same.
2. **Sub-structure search.** Returns whether or not the query molecule is a substructure of the target molecule. If the molecules are represented as 2D graph where atoms are indices and bonds between them are vertices. Then sub-structure search can be approached as subgraph isomorphism problem, where the task is to find out whether of two given graphs A and B, A contains a subgraph that is isomorphic to B.
3. **Similarity search.** Returns whether or not the Tanimoto similarity between two molecules (fingerprints) exceeds the cut-off value.

DATABASE OVERVIEW

Database content

ConjuPepDB contains information of >1600 unique peptidic conjugates collected from 230 published research articles. It provides comprehensive information about a peptidic conjugate such as 2D and 3D models, molecular properties, compound identifier (CAS RN for both the conjugate and the small molecule alone), structural information including peptide sequence, covalent binding mode as well as conjugated small molecule structure, bibliography and other information (Figure 4). Data are organised in several fields such as chemical structure, chemical name, source, SMILES, InchiKey, molecular weight, molecular formula, CAS RN, application, type of conjugate and references. Additionally, various calculated properties such as log *P*, num-

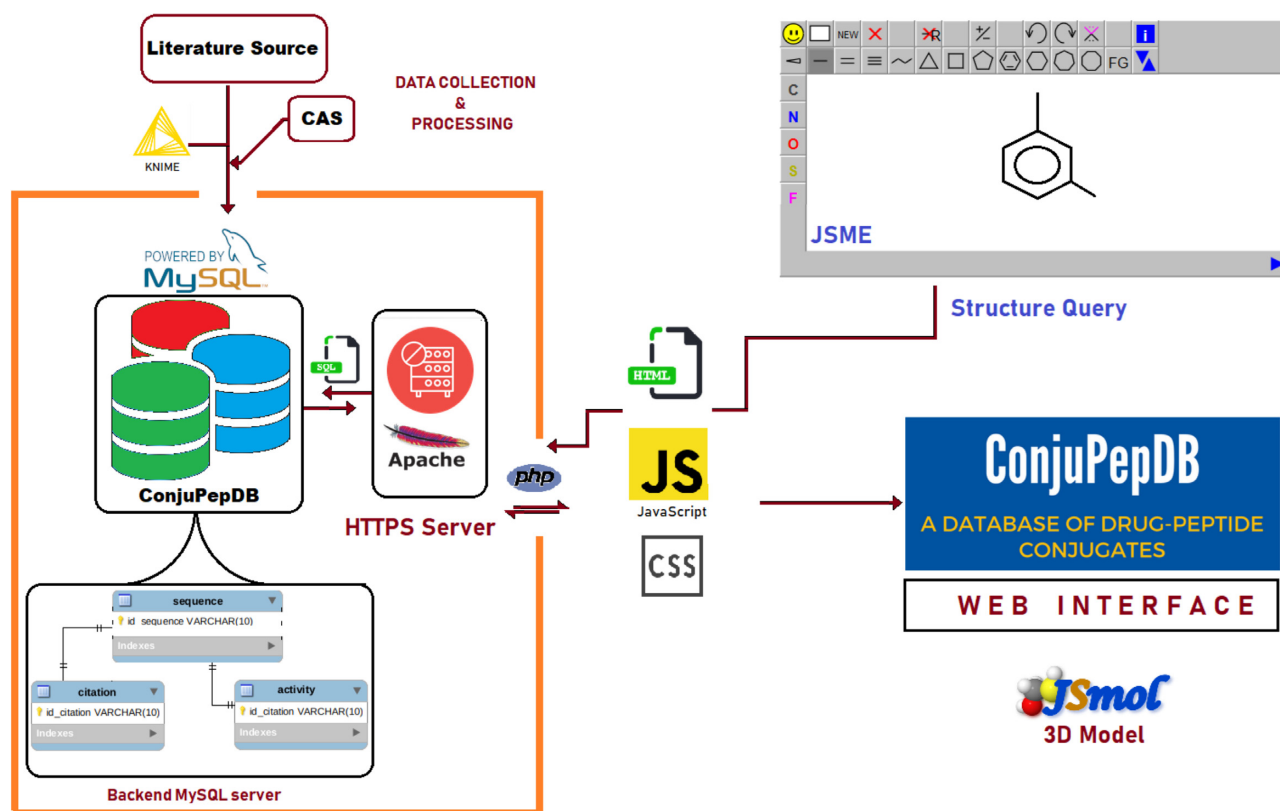


Figure 3. A schematic diagram of workflow of data collection and processing and layout of information retrieval in ConjuPepDB.

ber of H-bond donors and acceptors, rotatable bonds and surface area are also provided.

User interface layout

ConjuPepDB is designed to be intuitive and user-friendly with multiple ways to navigate the database. Options are provided to browse all conjugates, articles, structures and biological activities. The rich interface for querying allows the user to easily retrieve a specific conjugate from the database. Users can perform a simple search beside the complex and chemical fingerprint-based substructure search. The web-interface is designed to be responsive and compatible with devices of different screen sizes. A brief description of the main pages in ConjuPepDB is as follows.

- Home:** Main landing page, with a brief introduction and statistical information about the database.
- Browse Conjugates:** This page has an interactive table spread over multiple sub-pages displaying all conjugates in ConjuPepDB. Options to filter the table based year of publications are also provided (Figure 4A). Details of a single conjugate can be accessed by clicking on the ConjuPepDB ID, which takes the user to the single conjugate view page.
- Search:** This page has comprehensive options to perform simple, complex and substructure searches (Figure 4C). The simple search allows performing a search using various fields such as internal ID (ConjuPepDB ID), CAS RN, application, article title, author name and journal. Structure search, in turn, allows the users to draw the query molecule and search for entries in the ConjuPepDB.
- Single conjugate view:** This page displays both a 2D and an interactive 3D model and all relevant information associated with a single conjugate. The animation and representation of the 3D model can be controlled using the buttons provided below the model (Figure 4B). A schematic illustration of the conjugate is shown, where the sequence of the peptide, the type of the linker and the name of the small organic drug molecule is displayed. Data in this page are divided into six categories as follows. (i) Identification (chemical name, sequence, SMILES, InchiKey, molecular weight, molecular formula, source). (ii) External ID (CAS RN). (iii) Structural data, which include information about the method used for structure determination. (iv) Other information, namely, application, conjugate type, etc. (v) Calculated properties such as LogP, number of H-bond donor or acceptor, rotatable bond and polar surface area (PSA). (vi) Information related to citations.
- Browse article:** List of the articles from which data are included in ConjuPepDB with links to the full-text options is presented in this page. The title of the article, author names, journal, year and the number of conjugates in each article are shown in the table (Figure 4D).
- Feedback:** This page holds the details on how to give feedback and report errors. The templates for contributing new data into ConjuPepDB are also available here.

A Browse Conjugates
Showing all from all years

Showing 31 - 60 of 951 Entries --- page 2 of 32 --- First Prev 1 2 3 4 5 ... 30 31 32 Next

Last

Y Filters	ConjuPepDB ID	CAS RN	Peptide name	Small molecule	Application	Conjugate Type	Reference
Article year	cpd00031	2305943-24-0	lytic peptides	Pterico acid	Drug-resistant cancer therapy	amide	Dai, Yousun et al (2019) DOI
All time - 2019-2015 2014-2010 2009-2005 2004-1999	cpd00032	2305943-23-9	lytic peptides	Pterico acid	Drug-resistant cancer therapy	amide	Dai, Yousun et al (2019) DOI
Clear filters	cpd00033	2305943-22-8	lytic peptides	Pterico acid	Drug-resistant cancer therapy	amide	Dai, Yousun et al (2019) DOI
	cpd00034	2305943-21-7	lytic peptides	Fluorescein	Drug-resistant cancer therapy	amide	Dai, Yousun et al (2019) DOI

B View Conjugates's details
ConjuPepDB ID: cpd00002

Chemical diagram 3D viewer
3D model is for representation purpose only.

CPK Stack B3S Hide H Show H

C Search
Search in ConjuPepDB

Simple search Structure search

Free text search
Search ConjuPepDB. Search

Following fields are searched:
ConjuPepDB ID, peptide sequence (one letter), CAS RN, Inchi key, application, peptide name, small molecule name, and conjugate types.

Use the following characters to construct logical search query:
+ stands for AND
- stands for NOT
[no operator] implies OR
The asterisk serves as the truncation (or wildcard) operator. It should be appended to the word to be affected. Words match if they begin with the word preceding the * operator.

Example: 1. Search query 'lytic+anticoor' will fetch the entries of type lytic and anticoor application.
2. Search query 'lytic+anticoor-amide' will fetch all the conjugates with anticoor application while excluding type amide.

See details of Boolean Full-Text Searches

Simple search
Select a keyword to search, then type the term you want to search in the below input, and press the search button.

Keywords: ConjuPepDB ID CAS RN Application

Search ConjuPepDB. Search

Example: ConjuPepDB ID: cpd00017, CAS RN: 2375240-36-9, Application: Antimicrobial

Article search
Select a keyword to search, then type the term you want to search in the below input, and press the search button.

Keywords: Title Authors Journal Year

Search article in ConjuPepDB. Search

Example: Title: N-Methyl-N-phenylvinylsulfonamides for Cysteine-Selective Conjugation, Author: Ahmed, Journal: ACS Chemical Biology, Year: 2010

D Browse articles
Browse conjugates by articles

Article ID A/Y	Title A/Y	Authors A/Y	Journal	Year A/Y	DOI	Records
cpdar0001	Accumulation of arginine-rich cell-penetrating peptides in tumors and the potential for anticancer drug delivery in vivo.	Nakase, Rukiho; Konehi, Yasuko; Ueda, Masaaki; Saji, Hideo; Fukui, Shiroh	Journal of controlled release: official journal of the Controlled Release Society	2012	DOI	1
cpdar0005	Design and biological evaluation of cell-penetrating peptide-doxorubicin conjugates as prodrugs.	Naezollahi Shirazi, Amir; Twari, Rakesh; Chhikara, Bhupender S; Mandal, Dindyal; Parang, Keykavous	Molecular pharmaceutics	2013	DOI	2
cpdar0012	Controlled release of free doxorubicin from peptide-drug conjugates by drug loading.	Chen, Zhipeng; Zhang, Pengcheng; Cheetham, Andrew G; Moon, Jae Hyon; Moxley, James W Jr; Lin, Yuan; Cui, Honggang	Journal of controlled release: official journal of the Controlled Release Society	2014	DOI	3
cpdar0014	Cell penetrating peptides as efficient nanocarriers for delivery of antifungal compound, natamycin for the treatment of fungal keratitis.	Jain, Aastha; Shah, Sushmita G; Chugh, Archana	Pharmaceutical research	2015	DOI	1
cpdar0015	Activable Cell-Penetrating Peptide Conjugated Prodrug for Tumor Targeted Drug Delivery.	Cheng, Hong; Zhu, Jing-Yi; Xu, Xiao-Ding; Qiu, Wen-Xiu; Lei, Qi; Han, Kai; Cheng, Yin-Jia; Zhang, Xian-Zhang	ACS Applied Materials & Interfaces	2015	DOI	1
cpdar0016	Dual-targeting hybrid peptide-conjugated doxorubicin for drug resistance reversal in breast cancer.	Sheng, Yuan; You, Yiwen; Chen, Yun	International Journal of Pharmaceutics	2016	DOI	1
cpdar0020	BH3 mimetics derived from Bim-BH3 domain core region show PTP1B inhibitory activity.	Lu, Xiao; Wu, Lijuan; Liu, Xiaohun; Wang, Shulin; Zhang, Chuanliang	Bioorganic and Medicinal Chemistry Letters	2019	DOI	1

E Article's details
Title
Design and biological evaluation of cell-penetrating peptide-doxorubicin conjugates as prodrugs.

Authors: Naezollahi Shirazi, Amir; Twari, Rakesh; Chhikara, Bhupender S; Mandal, Dindyal; Parang, Keykavous; Molecular pharmaceutics; (2013); 10.1021/mp3004034

Abstract
Doxorubicin (Dox) is a hydrophilic anticancer drug that has short retention time due to the efficient efflux in some cancer cells (e.g., ovarian adenocarcinoma SK-OV-3). Cyclic [WRW(4)] and the corresponding linear peptide (RW(4)) were conjugated with Dox through an appropriate linker to afford cyclic [WRW(4)]-Dox and linear (RW(4))-Dox conjugates to enhance the cellular uptake and cellular retention of the parent drug for sustained anticancer activity. Comparative antiproliferative assays between covalent cyclic [WRW(4)]-Dox and linear (RW(4))-Dox and the corresponding noncovalent physical mixtures of the peptides and Dox were performed. Cyclic [WRW(4)]-Dox inhibited the cell proliferation of human leukemia (CCRF-CEM) (62.7%), ovarian adenocarcinoma (SK-OV-3) (51.74%), colorectal carcinoma (HCT-116) (50.67%), and breast carcinoma (MDA-MB-468) (60.79%) cells at a concentration of 1.7 μM after 72-120 h of incubation. Cyclic [WRW(4)]-Dox exhibited higher antiproliferative activity than linear (RW(4))-Dox in all cancer cells with the highest activity observed after 72 h. Flow cytometry analyses showed 3.6-fold higher cellular uptake of cyclic [WRW(4)]-Dox than Dox alone in SK-OV-3 cells after 24 h incubation. The cellular hydrolysis study showed that 99% of cyclic [WRW(4)]-Dox was hydrolyzed intracellularly within 72 h and released Dox. These data suggest that cyclic [WRW(4)]-Dox can be used as a potential prodrug for improving the cellular delivery and retention of Dox.

Figure 4. Overview of different ConjuPepDB webpages. Layout of: (A) browse conjugates table along with the filters, (B) structures and related data of a conjugate entry, (C) search options in ConjuPepDB, (D) browse article table, (E) single article page.

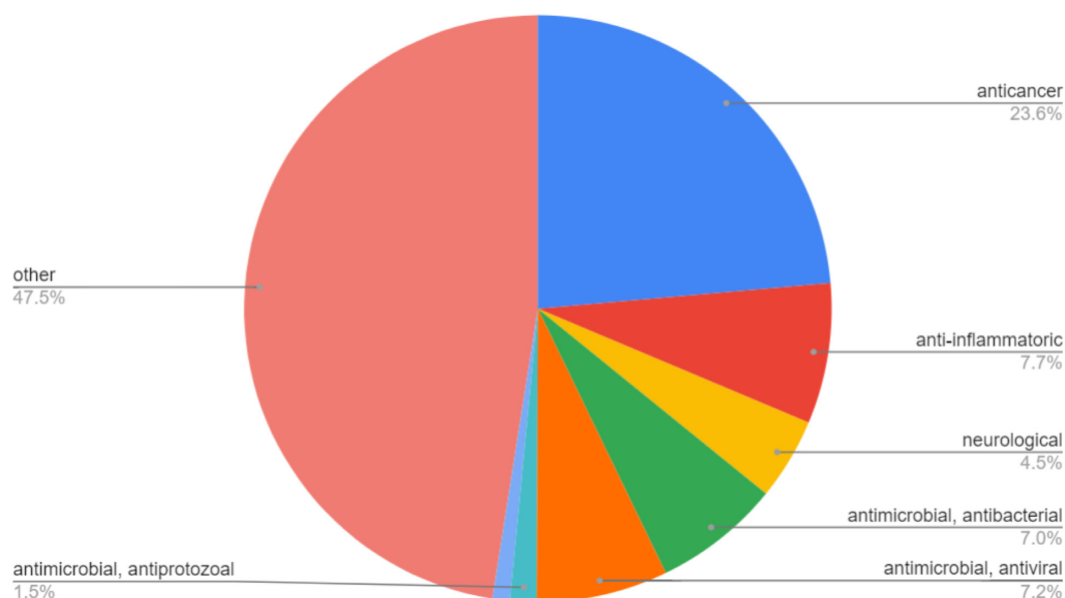


Figure 5. The distribution of applications found for the database entries.

Data submitted by users would be reviewed by the ConjuPepDB team before uploading.

Analysis of conjugate applications

The biomedical applications are included and their distribution in the ConjuPepDB are shown in Figure 5. The most common types of application are anticancer (372) and antimicrobial (262) utilizations.

Some anticancer drugs are poorly soluble in water or they are toxic. These disadvantageous properties can be remedied with conjugation of a small peptide (31). Other conjugates are designed as ‘prodrugs’ activated by an enzyme produced by the tumor cells by cleaving the peptide part thereby activating the compound (23). The application of tumor specific peptides made the targeted therapy available: the Human epidermal growth factor (HER2) receptor (32) and the gonadotropin-releasing hormone (GnRH) receptors are typical examples for these (33). With the addition of a cell-penetrating peptide the anticancer drug can be more effectively cumulated in the tumor cells (34).

The antibacterial conjugates (110) are the second most numerous group of conjugates analysed (35–37). In most cases, these conjugates are useful to increase the selectivity or the activity (or decrease the resistance) of the compounds compared to their unconjugated forms. As an example, the antimycobacterial activity of peptide conjugate of pyridopyrimidine derivatives were found to be more effective and less toxic in a series of *in vivo* and *in vitro* testing than the original small organic drug molecule (38). Antiprotozoal (23 conjugates, mostly antimalarial (11)), antiviral (114 conjugates, for e.g. HIV-1, Hepatitis-C and Herpes simplex virus) (9,39,40) and antifungal (15 conjugates) (10) also occur in the literature, albeit in a smaller number.

A large number (122 conjugates) of anti-inflammatory drug-conjugates were also found, e.g. for the treatment of rheumatoid arthritis (41), while conjugates targeting spe-

cific tissues (such as the bone with ibuprofen) have been also investigated (42).

Diseases of the central nervous system, mainly Alzheimer’s disease (14) also seems to be a promising area of research (71 conjugates).

Although drug-conjugates with biomedical action were the main focus of our literature search, a huge number of other conjugates (749) were found without any specific pharmacological properties. These were mostly related to various synthetic efforts to invent and improve the conjugation techniques without any biological tests (43–45). Various other basically rare biomedical applications, including facilitated blood–brain barrier transport (46), biotechnological applications like DNA cleavage (47), radiolabeling (48), imaging (49) or chelation were published too (50). For the sake of the database, these are valuable entries, because these appreciate the more diverse applicability of peptide conjugates and also provide useful insight to synthetic solutions for interested organic chemists.

Examples

Brief examples of a few search options provided in ConjuPepDB are described below:

1. **Free text search:** A free text based search has been implemented which allows users to perform keyword-based search on various fields of the database. Database fields included in this search are ConjuPepDB ID, peptide sequence (one letter code), CAS RN, InChikey, application, peptide name, small molecule name, and conjugate types. Logical search query can be performed by using logical operators such as ‘+’ for AND, ‘-’ for NOT and ‘no operator’ for OR. For example, if we search using ‘sulfide + anticancer’ as the search term, it will return 387 hits of conjugates of type sulfide with anticancer properties (Figure 6). On the other hand, search query

Simple search

Free text search ?

Following fields are searched.
ConjuPepDB ID, peptide sequence (one letter), CAS RN, inchi key, application, peptide name, small molecule name, and conjugate types.

Use the following characters to construct logical search query.
+ stands for AND
- stands for NOT
[no operator] implies OR
The asterisk serves as the truncation (or wildcard) operator. It should be appended to the word to be affected. Words match if they begin with the word preceding the * operator.

Example: 1. Search query 'sulfide+anticancer' will fetch the entries of type sulfide and anticancer applicaton.
2. Search query 'anticancer-amide' will fetch all the conjugates with anticancer application while excluding type amide.

A**Free text search results...**

Searched term 'sulfide+anticancer' in free text search.

Total 118 hits found...

#	ConjuPepDB ID	CAS RN	Peptide name	Small molecule	Type	Reference
1	cpd00009	2000174-85-4		doxorubicin	sulfide	cpdart0016
2	cpd00361	1541068-43-2	NoxaBH3	6,6'-Bis-bromomethyl-[3,3']bipyridine	sulfide	cpdart0068
3	cpd00526	1427500-02-4	TAT-analogue	Doxorubicin	sulfide	cpdart0082
4	cpd00635	1427354-59-3	activable cell-penetrating peptide (ACCP)	Doxorubicin	sulfide	cpdart0100
5	cpd00735	1350616-79-3	cell-penetrating octaarginine and lung-targeting peptide	Pemetrexed	sulfide	cpdart0115
6	cpd00736	1350616-78-2	octaarginine	Pemetrexed	sulfide	cpdart0115
7	cpd00737	1350616-77-1	lung-targeting peptide	Pemetrexed	sulfide	cpdart0115
8	cpd00738	1350616-76-0	cell-penetrating octaarginine and lung-targeting peptide	Pemetrexed	sulfide	cpdart0115
9	cpd00739	1350616-75-9	octaarginine	Pemetrexed	sulfide	cpdart0115
10	cpd00740	1350616-74-8	lung-targeting peptide	Pemetrexed	sulfide	cpdart0115
11	cpd00008	1801419-50-0		doxorubicin	sulfide	cpdart0015
12	cpd00001	143413-47-2	R8			cpdart0001
13	cpd00002	1417821-47-6		doxorubicin	amide	cpdart0005

B

Figure 6. (A) The interface for free text based search in ConjuPepDB. (B) The search results of the query. A total of 118 hits were obtained for query 'sulfide+anticancer'.

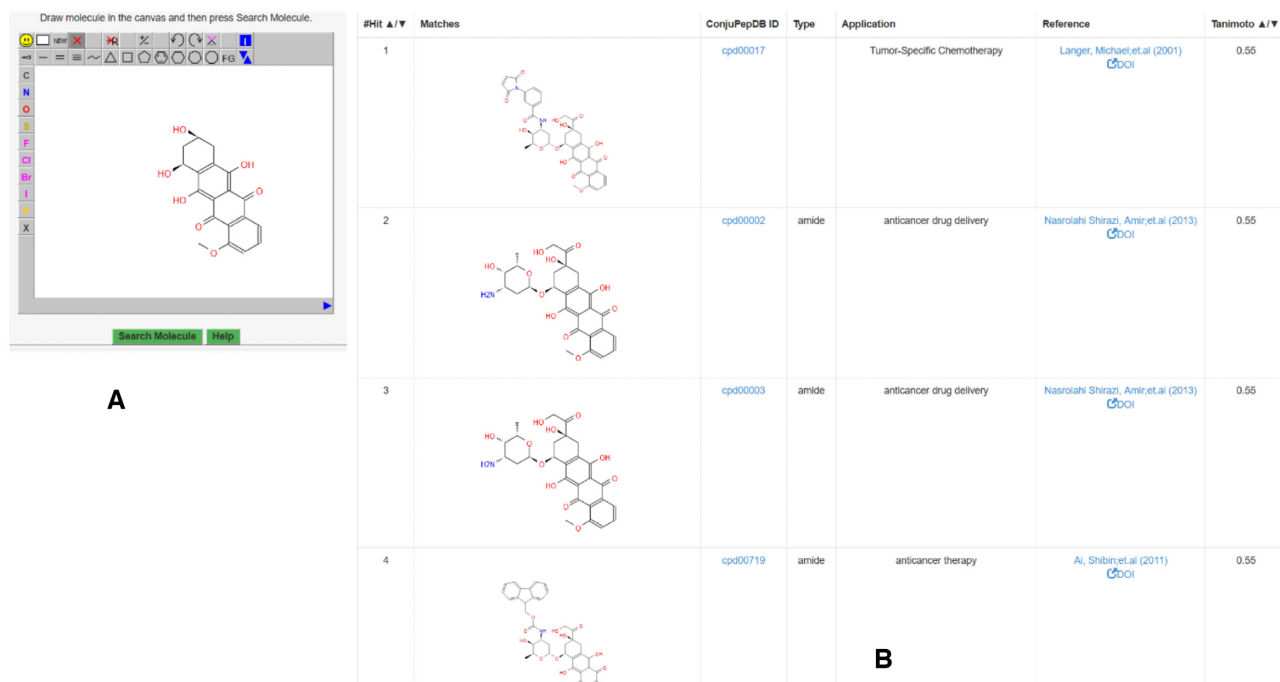


Figure 7. (A) The query structure drawn in JSME editor. (B) Hits of the chemical structure search obtained with Tanimoto score.

‘anticancer-amide’ will collect all the conjugates with anticancer application while excluding type amide. Additional information of this topic is available at https://conjupepdb.ttk.hu/help.php#Simple_Search.

- Simple Search:** Specific entries in ConjuPepDB can also be searched using the Internal ID of the database (ConjuPepDB ID), CAS RN, author name, title of the research article, journal name, as well as year of publication.
- Structure search:** We have also implemented a chemical search functionality. Here users can sketch the query structure in a graphical interface and run the structure search. The search is performed on the small molecule part of the peptide–drug conjugate only. With this a user can easily find the relevant conjugates with a drug like chemical scaffold of interest. As an example, we use the query molecule shown in Figure 7A, which is represented in smile format as:

```
'COc4cccc5c(=O)c3c(O)c2C[C@@](O)(C(=O)CO)C[C@H](O)[C@H]1C[C@H](N)[C@H](O)[C@H](C)O1)c2c(O)c3c(=O)c45'
```

Beside sketching it from scratch it is also possible to paste the SMILES notation of our query structure into the canvas. Readers are encouraged to look for more details on this option in the ConjuPepDB help page which can be accessed at https://conjupepdb.ttk.hu/help.php#Sub-structure_Search. After drawing the query molecule and selecting the type of chemical search, press the ‘SEARCH’ button, which will take you to the result page once the search is finished. Figure 7B shows the output of your structure search, where the hits are sorted using Tanimoto distance

between the query and the hit molecule. Molecules with higher similarity have higher Tanimoto coefficients, closer to 1, while less similar ones have lower values, closer to 0. In addition, users can also search the database by exact structure or sub-structure search.

SUMMARY AND OUTLOOK

In summary, we created the ConjuPepDB which is a publicly available database comprising information on peptide–drug conjugates. Data were collected manually and treated to summarise and present the relevant information of this important and rapidly developing field of potential drug molecules. It comprises basic information on each compound such as the structure of the conjugate, the CAS RN, the name of the peptide and the small organic drug molecule. Furthermore, the entries also include information on the application of the conjugate, and the type of the chemical linkage that was chosen for covalent coupling of peptide and the drug molecule. ConjuPepDB contains 1645 entries with biomedical application collected from 238 research papers from 95 different scientific journals. According to our knowledge, this is the first freely available comprehensive database on peptide–drug conjugates. We believe that this DB might disseminate the importance of this class of organic compounds with significant future biomedical potential for industrial applications.

DATA AVAILABILITY

ConjuPepDB is an open source collaborative initiative available in the TTK Lendület Artificial Transporter Research Group repository (<https://conjupepdb.ttk.hu/>).

ACKNOWLEDGEMENTS

The authors are grateful to Dr Ákos Bencsura for technical help and discussions.

FUNDING

Hungarian Academy of Sciences [LP2016-2, LP2017-8]; National Competitiveness and Excellence Program [NVKP_16-1-2016-0007]; BIONANO_GINOP-2.3.2-15-2016-00017b; Project no. 2018-1.2.1-NKP-2018-00005 has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the 2018-1.2.1-NKP funding scheme. Funding for open access charge: Lendület [LP2017-8].

Conflict of interest statement. None declared.

REFERENCES

- Vrettos, E.I., Mezo, G. and Tzakos, A.G. (2018) On the design principles of peptide–drug conjugates for targeted drug delivery to the malignant tumor site. *Beilstein J. Org. Chem.*, **14**, 930–954.
- He, R., Finan, B., Mayer, J.P. and DiMarchi, R.D. (2019) Peptide conjugates with small molecules designed to enhance efficacy and safety. *Molecules*, **24**, 1855.
- Wang, M., Rakesh, K.P., Leng, J., Fang, W.Y., Ravindar, L., Channe Gowda, D. and Qin, H.L. (2018) Amino acids/peptides conjugated heterocycles: a tool for the recent development of novel therapeutic agents. *Bioorg. Chem.*, **76**, 113–129.
- Polgar, L., Lajko, E., Soos, P., Lang, O., Manea, M., Merkely, B., Mezo, G. and Kohidai, L. (2018) Drug targeting to decrease cardiotoxicity - determination of the cytotoxic effect of GnRH-based conjugates containing doxorubicin, daunorubicin and methotrexate on human cardiomyocytes and endothelial cells. *Beilstein J. Org. Chem.*, **14**, 1583–1594.
- Lajko, E., Spring, S., Hegedus, R., Biri-Kovacs, B., Ingebrandt, S., Mezo, G. and Kohidai, L. (2018) Comparative cell biological study of in vitro antitumor and antimetastatic activity on melanoma cells of GnRH-III-containing conjugates modified with short-chain fatty acids. *Beilstein J. Org. Chem.*, **14**, 2495–2509.
- Yamada, R., Kostova, M.B., Anchoori, R.K., Xu, S., Neamati, N. and Khan, S.R. (2010) Biological evaluation of paclitaxel-peptide conjugates as a model for MMP2-targeted drug delivery. *Cancer Biol. Ther.*, **9**, 192–203.
- Demeter, O., Kormos, A., Koehler, C., Mezo, G., Nemeth, K., Kozma, E., Takacs, L.B., Lemke, E.A. and Kele, P. (2017) Bisazide cyanine dyes as fluorogenic probes for Bis-Cyclooctynylated peptide tags and as fluorogenic cross-linkers of cyclooctynylated proteins. *Bioconjug. Chem.*, **28**, 1552–1559.
- Baranyai, Z., Kratky, M., Vosatka, R., Szabo, E., Senoner, Z., David, S., Stolarikova, J., Vinsova, J. and Bosze, S. (2017) In vitro biological evaluation of new antimycobacterial salicylanilide-tuftsin conjugates. *Eur. J. Med. Chem.*, **133**, 152–173.
- Diez-Torrubia, A., Cabrera, S., de Castro, S., Garcia-Aparicio, C., Mulder, G., De Meester, I., Camarasa, M.J., Balzarini, J. and Velazquez, S. (2013) Novel water-soluble prodrugs of acyclovir cleavable by the dipeptidyl-peptidase IV (DPP IV/CD26) enzyme. *Eur. J. Med. Chem.*, **70**, 456–468.
- Ferreira, S.Z., Carneiro, H.C., Lara, H.A., Alves, R.B., Resende, J.M., Oliveira, H.M., Silva, L.M., Santos, D.A. and Freitas, R.P. (2015) Synthesis of a new peptide-coumarin conjugate: a potential agent against cryptococcosis. *ACS Med. Chem. Lett.*, **6**, 271–275.
- Perez, B.C., Teixeira, C., Figueiras, M., Gut, J., Rosenthal, P.J., Gomes, J.R. and Gomes, P. (2012) Novel cinnamic acid/4-aminoquinoline conjugates bearing non-proteinogenic amino acids: towards the development of potential dual action antimalarials. *Eur. J. Med. Chem.*, **54**, 887–899.
- Shokri, B., Zarghi, A., Shahhoseini, S., Mohammadi, R. and Kobarfard, F. (2019) Design, synthesis and biological evaluation of peptide-NSAID conjugates for targeted cancer therapy. *Arch. Pharm. (Weinheim)*, **352**, e1800379.
- Homma, A., Sato, H., Tamura, T., Okamachi, A., Emura, T., Ishizawa, T., Kato, T., Matsuura, T., Sato, S., Higuchi, Y. *et al.* (2010) Synthesis and optimization of hyaluronic acid-methotrexate conjugates to maximize benefit in the treatment of osteoarthritis. *Bioorg. Med. Chem.*, **18**, 1062–1075.
- Schieb, H., Weidlich, S., Schlechtingen, G., Linning, P., Jennings, G., Gruner, M., Wiltfang, J., Klafki, H.W. and Knolker, H.J. (2010) Structural design, solid-phase synthesis and activity of membrane-anchored β -secretase inhibitors on A β generation from wild-type and Swedish-mutant APP. *Chem. Eur. J.*, **16**, 14412–14423.
- Tang, Q., Cao, B., Wu, H. and Cheng, G. (2013) Cholesterol-peptide hybrids to form liposome-like vesicles for gene delivery. *PLoS One*, **8**, e54460.
- Miller, D.S., Scambia, G., Bondarenko, I., Westermann, A.M., Oaknin, A., Oza, A.M., Lisyanskaya, A.S., Vergote, I., Wenham, R.M., Temkin, S.M. *et al.* (2018) ZoptEC: Phase III randomized controlled study comparing zoletarelin with doxorubicin as second line therapy for locally advanced, recurrent, or metastatic endometrial cancer (NCT01767155). *J. Clin. Onc.*, **36**, 5503.
- Engel, J. B., Tinneberg, H.R., Rick, F.G., Berkes, E. and Schally, A.V. (2016) Targeting of peptide cytotoxins to LHRH receptors for treatment of cancer. *Curr. Drug Targets*, **17**, 488–494.
- Langer, C.J., O'Byrne, K.J., Socinski, M.A., Mikhailov, S.M., Leśniewski-Kmak, K., Smakal, M., Ciuleanu, T.E., Orlov, S.V., Dediu, M., Heigener, D. *et al.* (2008) Phase III trial comparing paclitaxel poliglumex (CT-2103, PPX) in combination with carboplatin versus standard paclitaxel and carboplatin in the treatment of PS 2 patients with chemotherapy-naïve advanced non-small cell lung cancer. *J. Thorac. Oncol.*, **3**, 623–630.
- Graybill, W.S. and Coleman, R.L. (2014) Vintafolide: a novel targeted agent for epithelial ovarian cancer. *Future Oncol.*, **10**, 541–548.
- Zhang, P., Cheetham, A.G., Lock, L.L. and Cui, H. (2013) Cellular uptake and cytotoxicity of peptide–drug conjugates regulated by conjugation site. *Bioconjug. Chem.*, **24**, 604–613.
- Parvatkar, P., Kato, N., Uesugi, M., Sato, S. and Ohkanda, J. (2015) Intracellular generation of a diterpene-peptide conjugate that inhibits 14-3-3-mediated interactions. *J. Am. Chem. Soc.*, **137**, 15624–15627.
- Bera, S., Zhanel, G.G. and Schweizer, F. (2010) Evaluation of amphiphilic aminoglycoside-peptide triazole conjugates as antibacterial agents. *Bioorg. Med. Chem. Lett.*, **20**, 3031–3035.
- Legigan, T., Clarhaut, J., Tranoy-Opalinski, I., Monvoisin, A., Renoux, B., Thomas, M., Le Pape, A., Lerondel, S. and Papot, S. (2012) The first generation of beta-galactosidase-responsive prodrugs designed for the selective treatment of solid tumors in prodrug monotherapy. *Angew. Chem. Int. Ed.*, **51**, 11606–11610.
- Borodkin, V.S., Schimpl, M., Gundogdu, M., Rafie, K., Dorfmueller, H.C., Robinson, D.A. and van Aalten, D.M. (2014) Substrate UDP-peptide conjugates as human O-GlcNAc transferase inhibitors. *Biochem. J.*, **457**, 497–502.
- Panda, S.S., Ibrahim, M.A., Kucukbay, H., Meyers, M.J., Sverdrup, F.M., El-Feky, S.A. and Katritzky, A.R. (2013) Synthesis and antimalarial bioassay of quinine - peptide conjugates. *Chem. Biol. Drug Des.*, **82**, 361–366.
- Wang, J., Hu, S., Mao, W., Xiang, J., Zhou, Z., Liu, X., Tang, J. and Shen, Y. (2019) Assemblies of peptide-cytotoxin conjugates for tumor-homing chemotherapy. *Adv. Funct. Mater.*, **29**, 1807446.
- Wang, Y., Cheethama, A.G., Angacian, G., Sua, H., Xie, L. and Cui, H. (2017) Peptide–drug conjugates as effective prodrug strategies for targeted delivery. *Adv. Drug Deliv. Rev.*, **110–111**, 112–126.
- Agrawal, P., Bhalla, S., Usmani, S.S., Singh, S., Chaudhary, K., Raghava, G.P. and Gautam, A. (2016) CPPsite 2.0: a repository of experimentally validated cell-penetrating peptides. *Nucleic Acids Res.*, **44**, D1098–D1103.
- Bienfait, B. and Ertl, P. (2013) JSME: a free molecule editor in JavaScript. *J. Cheminform.*, **5**, 24.
- Nizami, B., Bereczki-Szakal, D., Varro, N., El Battioui, K., Nagaraj, V.U., Szigartyo, I.C., Mandity, I. and Beke-Somfai, T. (2020) FoldamerDB: a database of peptidic foldamers. *Nucleic Acids Res.*, **48**, D1122–D1128.
- Goswami, D., Vitorino, H.A., Machini, M.T. and Esposito, B.P. (2015) Self-assembled penetratin-deferasirox micelles as potential carriers for hydrophobic drug delivery. *Biopolymers*, **104**, 712–719.

32. Biri-Kovacs,B., Adorjan,A., Szabo,I., Szeder,B., Bosze,S. and Mezo,G. (2020) Structure-activity relationship of HER2 receptor targeting peptide and its derivatives in targeted tumor therapy. *Biomolecules*, **10**, 183.
33. Szabo,I., Bosze,S., Orban,E., Sipos,E., Halmos,G., Kovacs,M. and Mezo,G. (2015) Comparative in vitro biological evaluation of daunorubicin containing GnRH-I and GnRH-II conjugates developed for tumor targeting. *J. Pept. Sci.*, **21**, 426–435.
34. Cheng,H., Zhu,J.Y., Xu,X.D., Qiu,W.X., Lei,Q., Han,K., Cheng,Y.J. and Zhang,X.Z. (2015) Activable cell-penetrating peptide conjugated prodrug for tumor targeted drug delivery. *ACS Appl. Mater. Interfaces*, **7**, 16061–16069.
35. Bera,S., Zhanel,G.G. and Schweizer,F. (2011) Synthesis and antibacterial activity of amphiphilic lysine-ligated neomycin B conjugates. *Carbohydr. Res.*, **346**, 560–568.
36. Brezden,A., Mohamed,M.F., Nepal,M., Harwood,J.S., Kuriakose,J., Seleem,M.N. and Chmielewski,J. (2016) Dual targeting of intracellular pathogenic bacteria with a cleavable conjugate of kanamycin and an antibacterial cell-penetrating peptide. *J. Am. Chem. Soc.*, **138**, 10945–10949.
37. Rakesh,K.P., Ramesh,S., Shivakumar and Gowda,D.C. (2018) Effect of low charge and high hydrophobicity on antimicrobial activity of the quinazolinone-peptide conjugates. *Russ. J. Bioorg. Chem.*, **44**, 158–164.
38. Horvati,K., Bacsá,B., Szabo,N., Fodor,K., Balka,G., Rusvai,M., Kiss,E., Mezo,G., Grolmusz,V., Vertessy,B. *et al.* (2015) Antimycobacterial activity of peptide conjugate of pyridopyrimidine derivative against *Mycobacterium tuberculosis* in a series of in vitro and in vivo models. *Tuberculosis (Edinb.)*, **95**, S207–S211.
39. Wang,C., Shi,W., Cai,L., Lu,L., Wang,Q., Zhang,T., Li,J., Zhang,Z., Wang,K., Xu,L. *et al.* (2013) Design, synthesis, and biological evaluation of highly potent small molecule-peptide conjugates as new HIV-1 fusion inhibitors. *J. Med. Chem.*, **56**, 2527–2539.
40. Nitsche,C., Schreier,V.N., Behnam,M.A., Kumar,A., Bartenschlager,R. and Klein,C.D. (2013) Thiazolidinone-peptide hybrids as dengue virus protease inhibitors with antiviral activity in cell culture. *J. Med. Chem.*, **56**, 8389–8403.
41. Conda-Sheridan,M., Lee,S.S., Preslar,A.T. and Stupp,S.I. (2014) Esterase-activated release of naproxen from supramolecular nanofibres. *Chem. Commun.*, **50**, 13757–13760.
42. Zhao,Y., He,D., Ma,L. and Guo,L. (2015) Synthesis and preliminary evaluation of novel bone-targeting NSAIDs Prodrugs based on glutamic acid oligopeptides. *Lett. Drug Des. Discov.*, **12**, 585–590.
43. Slosarczyk,A.T. and Baltzer,L. (2012) Efficient formation of heterodimers from peptides and proteins using unsymmetrical polyfluorophenyl esters of dicarboxylic acids. *J. Pept. Sci.*, **18**, 261–269.
44. Nenajdenko,V., Osipov,S., Sokolova,N., Vorobyeva,D. and Vasilyeva,T. (2011) Synthesis of α -trifluoromethyl- α -hydroxy acid-peptide conjugates via click chemistry. *Synthesis*, **44**, 130–136.
45. van der Heden van Noort,G.J., van der Horst,M.G., Overkleeft,H.S., van der Marel,G.A. and Filippov,D.V. (2010) Synthesis of mono-ADP-ribosylated oligopeptides using ribosylated amino acid building blocks. *J. Am. Chem. Soc.*, **132**, 5236–5240.
46. Wangler,C., Chowdhury,S., Hofner,G., Djurova,P., Purisima,E.O., Bartenstein,P., Wangler,B., Fricker,G., Wanner,K.T. and Schirmacher,R. (2014) Shuttle-cargo fusion molecules of transport peptides and the hD2/3 receptor antagonist fallypride: a feasible approach to preserve ligand-receptor binding? *J. Med. Chem.*, **57**, 4368–4381.
47. Li,Z.F., Chen,H.L., Zhang,L.J. and Lu,Z.L. (2012) Synthesis of [12]janeN3-dipeptide conjugates as metal-free DNA nucleases. *Bioorg. Med. Chem. Lett.*, **22**, 2303–2307.
48. Jamous,M., Haberkorn,U. and Mier,W. (2012) DOTA-tris(OPp ester) as a bifunctional prochelator for the preparation of DOTA-peptide conjugates. *Tetrahedron Lett.*, **53**, 6810–6814.
49. Lee,S., Kang,S.W., Ryu,J.H., Na,J.H., Lee,D.E., Han,S.J., Kang,C.M., Choe,Y.S., Lee,K.C., Leary,J.F. *et al.* (2014) Tumor-homing glycol chitosan-based optical/PET dual imaging nanoprobe for cancer diagnosis. *Bioconjug. Chem.*, **25**, 601–610.
50. Sengupta,S., Krishnan,M.A., Dudhe,P., Reddy,R.B., Giri,B., Chattopadhyay,S. and Chelvam,V. (2018) Novel solid-phase strategy for the synthesis of ligand-targeted fluorescent-labelled chelating peptide conjugates as a theranostic tool for cancer. *Beilstein J. Org. Chem.*, **14**, 2665–2679.