ELSEVIER

THPdb2: compilation of FDA approved therapeutic peptides and proteins

Shipra Jain [1,#], Srijanee Gupta [1,#], Sumeet Patiyal [2], Gajendra P.S. Raghava [1,*]

[1] Department of Computational Biology, Indraprastha Institute of Information Technology, Okhla Phase 3, New Delhi 110020, India
[2] Cancer and Data Science Laboratory (CDSL), National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

During the past 20 years, there has been a significant increase in the number of protein-based drugs approved by the US Food and Drug Administration (FDA). This paper presents THPdb2, an updated version of the THPdb database, which holds information about all types of protein-based drugs, including peptides, antibodies, and biosimilar proteins. THPdb2 contains a total of 6,385 entries, providing comprehensive information about 894 FDA-approved therapeutic proteins, including 354 monoclonal antibodies and 85 peptides or polypeptides. Each entry includes the name of therapeutic molecule, the amino acid sequence, physical and chemical properties, and route of drug administration. The therapeutic molecules that are included in the database target a wide range of biological molecules, such as receptors, factors, and proteins, and have been approved for the treatment of various diseases, including cancers, infectious diseases, and immune disorders.

Keywords: Protein-based drugs; Monoclonal antibodies; Biosimilar proteins; Peptide therapeutics; Knowledgebase; FDA-approved drugs

Introduction

Proteins are an integral part of biological processes, which is why they are often referred to as the servants of living organisms. In the post-genomic era, protein databases are growing at an exponential rate due to advancements in sequencing technology. The latest release of UniProt contains around 248 million protein sequences. One of the major challenges in the current era is to annotate these proteins in order to understand their structure, function, and therapeutic potential. In this study, our major emphasis is on the therapeutic applications of proteins. Proteins and peptides have emerged as both therapeutics and diagnostic agents due to their potency, safety and selectivity, reshaping the pharmaceutical industry drastically.[p1] Current developments in the field of therapeutic proteins include improvements in incremental efficacy, safety, quality, and cost, and are based on research into new targets.[p2],[p3],[p4] The estimated global revenue generated by protein/peptide therapeutics is expected to shoot up to US$ 566.82 billion by 2030. Similarly, the global revenue generated by monoclonal antibodies is estimated to be US$ 679.03 billion by 2033.[p5] In addition to these protein products, various structural modifications, such as the use of unnatural amino acids, mainchain modifications, and other novel substitutions, have significantly improved the stability and therapeutic potential of peptides.[p3] The success of these modifications has inspired significant advancements in proteomics and artificial intelligence that have impacted the discovery of novel proteins directly. There is a need to screen these novel proteins for their therapeutic potential.

In order to facilitate the scientific community in fully harnessing the potential of protein-based drugs, computational biologists are constructing databases of proteins and developing tools for annotating therapeutic proteins. Historically, numerous data-

* *Corresponding author. Raghava, G.P.S. (raghava@iiitd.ac.in)*
# These authors contributed equally.

https://doi.org/10.1016/j.drudis.2024.104047

bases have been established to hold detailed information about bioactive proteins that have therapeutic potential.[p6],[p7],[p8] These repositories encompass a broad spectrum of inhibitory peptides, including anticancer (CancerPPD[p9]), antimicrobial (CAMPR4,[p10] APD3,[p11] and AntiTbPdb[p12]), and antihypertensive (AHTPDB,[p13] and BIOPEP[p14]) peptides, among others. Likewise, databases dedicated to peptides have emerged for drug delivery (CPPsite2[p15] TumorHope[p16] and B3Pdb[p17]), vaccine design (BCIPEP,[p18] MHCBN4,[p19] and IEDB[p20]), toxicity prediction (Hemolytik[p21] and COMPARE[p22]), and the development of stable peptides within bodily fluids (PEPlife[p23]) Despite the vast number of proteins that captured by these databases, only a select few protein-based drugs have successfully navigated all phases of clinical trials and gained approval from the US Food and Drug Administration (FDA) for clinical application. Hence, there exists a pressing need to establish a knowledge base specifically for FDA-approved protein-based drugs, encompassing peptides, polypeptides, monoclonal antibodies, polyclonal antibodies, and biosimilar proteins.

The International ImMunoGeneTics Information System (IMGT) has been maintaining a database dedicated to monoclonal antibodies, known as IMGT/mAb-DB, since 2010.[p24],[p25] IMGT/mAb-DB serves as a significant repository for therapeutic antibodies, with the current version containing 530 monoclonal antibodies for oncological use, of which 54 have received approval for clinical use from various regulatory agencies such as the FDA.[p26] To offer comprehensive information on all types of protein-based drugs, including peptides, antibodies, and biosimilar proteins, our research group undertook a systematic effort to establish the THPdb database of all FDA-approved therapeutic proteins.[p27] The original THPdb database encompasses 852 entries, providing detailed information on 239 protein-based drugs and their variants. In addition, it offers information on 59 monoclonal antibodies that have been approved by the FDA. THPdb has been extensively utilized and referenced by the scientific community, garnering approximately 490 citations to date. In the 7 years since the inception of THPdb, there has been a significant increase in the number of FDA-approved protein drugs. Consequently, there was a compelling need to develop an updated version of THPdb to enhance its service to the scientific community.

We introduce an updated version of THPdb, named THPdb2, which holds three times more information on therapeutic proteins than its predecessor. The data for THPdb2 were meticulously collected and curated from literature sources and publicly available databases such as DrugBank,[p28] PubChem,[p29] and UniProt.[p30] These therapeutic proteins have received approval for the treatment of a diverse array of diseases, spanning cancer, infectious diseases, immunological disorders, and metabolic disorders. THPdb2 offers a user-friendly web interface, which is freely accessible at https://webs.iiitd.edu.in/raghava/thpdb2/. Within this web server, various search modules are provided to empower users to explore datasets on the basis of diverse parameters, including disease area, physical attributes, category, administration route, target activity, and sequence length, among others. Subsequent sections will provide detailed insights into these features, highlighting the potential of THPdb2 as a valuable tool in the realm of therapeutic protein discovery and development.

Data collection and compilation

We have used a data collection approach similar to that used for THPdb.[p27] First, DrugBank[p28] was screened using the keywords "biotech drugs", along with "Approved" and "Investigational" filters, to extract the latest FDA-approved protein and peptide therapeutics data. This search yielded 831 biotechnology-based therapeutic protein and peptide drugs. Among these entries, only 387 had their peptide and protein sequences available in major repositories such as Uniprot,[p30] PubChem,[p29] and DrugBank.[p28] In addition, other public repositories were screened to extract comprehensive information on protein/peptide drugs, including their name, weight, chemical formula, isoelectric point, hydrophobicity, melting point (°C), half-life, description, mechanism, treated disease, pharmacodynamics, metabolism, toxicity, absorption, volume of distribution, clearance, categories, patents information, drug interactions and target. We also recorded drug information such as brand, formulation, dosage, side effects, and route of administration. Information for the 831 therapeutic proteins that was not available DrugBank was sought from PubMed and patent databases using keywords such as "therapeutic proteins", "therapeutic peptides", and the names of individual proteins and peptides. The information collected from these sources was compiled manually.

In order to capture all modifications made to approved drug molecules, THPdb2 includes entries for modified molecules as drug variants. Any changes or modifications to therapeutic proteins reported in THPdb are also appended to THPdb2; therefore, multiple entries are present in the new repository if a drug molecule has been modified in different ways or if it is being sold under different brand names. The final THPdb2 compilation includes the 239 therapeutic protein and peptide drugs that were included in THPdb,[p27] along with 655 new approved drugs.

Database architecture and organization

THPdb2 has been developed using the Apache HTTP server (version 2.4.7) and MySQL (version 5.5.62) to store and manage the data. HTML5, PHP5, CSS3 and JAVA script were deployed to create responsive front-end pages that are compatible with mobile phones, tablets and computers. MySQL was used as the back end of the database. A combined interface and common gateway were built using the PERL and PHP programming languages. The detailed architecture of THPdb2 is illustrated in Figure 1.

Web interface and modules

Search

Basic and advanced search modules have been developed in THPdb2. These modules enable users to search a given therapeutic protein or peptide based on terms such as Name, Molecular weight, Disease name, Category, Drug target, Brand, and Route of administration. In the basic search module, the users can customize the output based on the search query. The advanced search feature allows users to apply numerous queries simultane-

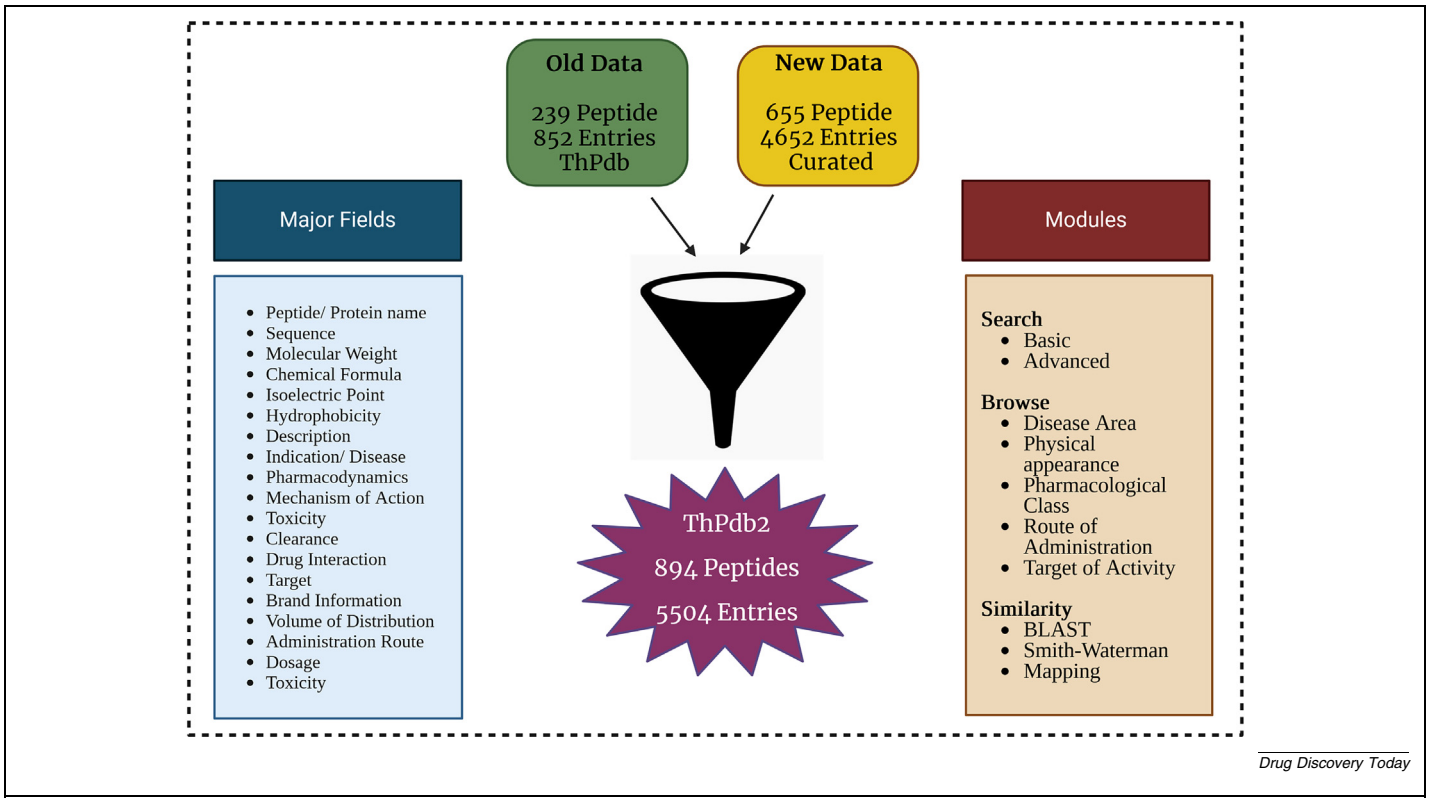

FIGURE 1
Schematic representation of the organization and structure of THPdb2.

ously using Boolean expressions (e.g. AND, NOT and OR) in order to obtain refined output. The THPdb2 module allows users to screen peptide and protein entries against desired physicochemical properties, and allows users to download the results in a.csv,.xls, or.pdf format.

Browse

In THPdb2, data can be browsed by several browse options: (i) Disease category—this module allows the user to choose the therapeutic peptides or proteins that are used against diseases in the selected category; (ii) Physical appearance—this module divides drugs in the dataset on the basis of which formulations (solution, powder, tablet, capsule, spray, or gel) are available in the market; (iii) Category/pharmacological class—this module allows the user to select a subset of the drugs in the dataset according to their pharmacological class, with classes including Antibodies, Enzymes, Antineoplastic agents, Hormones, and Immunosuppressive agents; (iv) Route of administration—this module helps the user to choose therapeutic proteins or peptides on the basis of their route of administration; (v) Target of activity—this module helps the user to select therapeutic peptides on the basis of their target type, which may be Receptors, Factors, Proteins, Enzymes, Hormones or Inhibitors; (vi) Sequence length – this module helps users to screen peptides, polypeptides and proteins on the basis of their sequence length. These browse options enable users to extract information on selected therapeutic peptides or proteins, which can be downloaded for each subcategory.

Similarity search

Our repository allows users to perform a similarity-based search for their input sequences. The Basic Local Alignment Search Tool[(p31)] (BLAST) and Smith–Waterman algorithms[(p32)] have been implemented in the webserver. This platform enables users to run sub-searches and super-searches using the Peptide Mapping module. When a sub-search is performed, a query peptide sequence is mapped against all peptides and proteins present in THPdb2 to find sequence matches; whereas a super-search returns all peptides in the database that are similar to the protein sequences given as a query. If no hit is found, then the Results page will be empty.

Data statistics

The THPdb2 database is an updated version of THPdb and provides the latest comprehensive information on FDA-approved therapeutic peptides and proteins. The first version of the database contains information on 239 peptides and proteins spread across 852 entries, and provides details about drug variants. In this updated version, we have incorporated modifications to the data captured in THPdb that have been observed since 2017. In this release, we have incorporated data on 239 old and 655 new (894 total) therapeutic protein- or peptide-based FDA-approved drugs. In order to capture data on drug variants, we have made multiple entries for the same peptide in the updated dataset. Please note that unique peptides have a unique therapeutic peptide ID (THPP_ID). Therefore, the earlier version

of data repository contains a total of 852 entries, whereas the updated version has a total of 6,385 entries.

The THPdb2 repository contains 579 therapeutic peptides or proteins that are used against cancer/tumors, 355 for infectious diseases, 350 for the treatment of immunological disease, 173 for cardiovascular disorders, 172 that are used to treat hormonal disorders, 160 used in neurological disorders, 126 against genetic disorders, 73 for respiratory disorders, 26 for hematological diseases, 25 for eye disorders, 12 that are effective in bone disorders, 9 that are used to treat metabolic disorders, and 3 that have been approved for malabsorption disorder, as well as 150 variants that are given as adjuncts (Figure 2A). When classified on the basis of their drug form, there were 209 solutions, 151 powders, 30 liquids, 19 in the form of cake, 17 resuspendable powders, 15 suspensions, 12 in tablet form, five in capsules, two used as sprays, two ointments and one gel (Figure 2B).

In order to understand the pharmacological classes of the therapeutic peptides and proteins, we have systematically extracted the data for various categories. This categorization shows 590 Antibodies, 382 Enzymes, 344 Antineoplastic agents, 238 Hormones, 193 Immunosuppressive agents, 38 Fibrinolytics, 34 Bone-related treatments, 25 fertility agents, 24 antithrombins, and 10 antidiabetic agents. We also divided the therapeutic peptides and proteins according to their route of administration and their drug targets. All of these statistics are summarized in Figure 3.

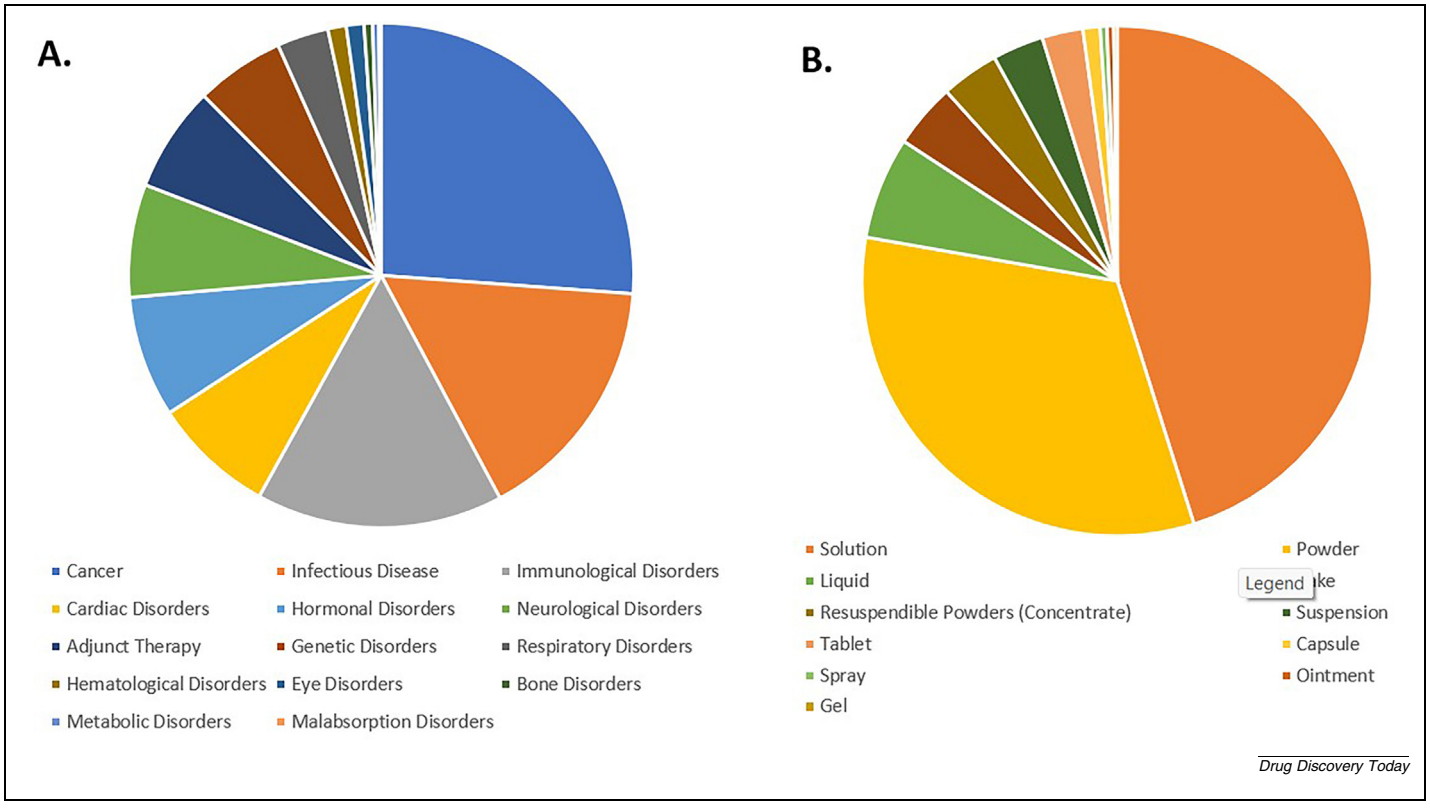

FIGURE 2

Categorization of therapeutic peptides and proteins based on A. disease type and B. drug form.

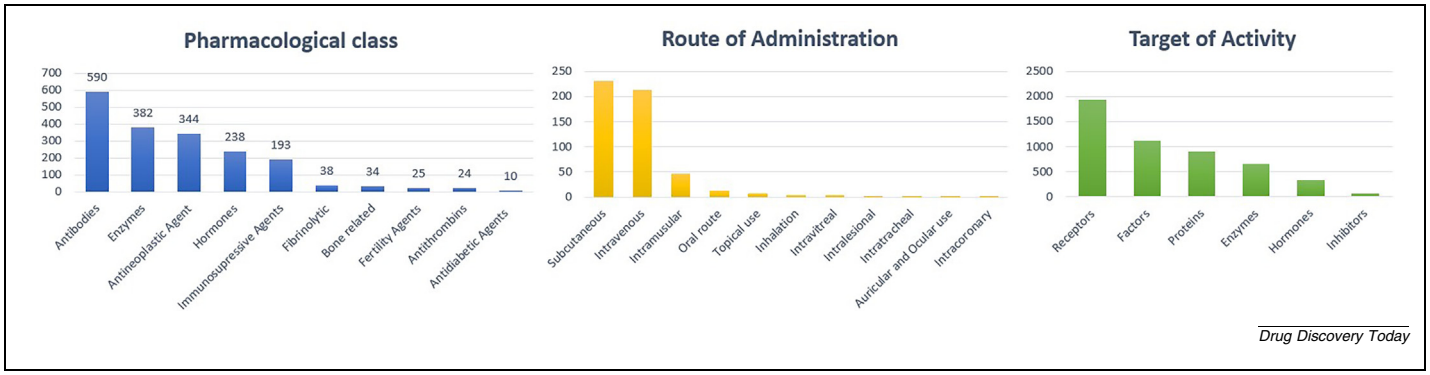

FIGURE 3

The distribution of therapeutic peptides and proteins based on (A) pharmacological class; (B) route of administration and (C) target of activity.

Discussion

In 1923, the first peptide-based drug, insulin, was utilized to treat diabetic patients who were unable to produce sufficient insulin or who had defective insulin.[p6],[p33] Subsequently, the advent of recombinant technology has provided a significant boost to the development of peptide-based drugs, including insulin, enabling the production of high-quality peptides on a large scale.[p34],[p35] Researchers have also explored the therapeutic potential of antibodies, which serve as effector molecules in the humoral immune response.[p36] In 1986, the first monoclonal antibody (mAb) drug was approved by the FDA for clinical applications. mAbs exhibit higher specificity, longer half-life, and fewer side effects when compared to traditional small-molecule-based drugs.

Broadly, therapeutic proteins can be divided into following three categories on the basis of their composition: (i) peptides/polypeptides, (ii) antibodies and (iii) non-antibody proteins. Peptides/polypeptides are those protein sequences that have fewer than 100 amino acids. Owing to technical advancements in the development of therapeutic peptides, more than 85 peptide-based drugs have now been approved for clinical applications and researchers' interest in the field of bioactive peptides continues to grow. In the past decade, hundreds of *in silico* methods and repositories have been developed to allow the prediction of therapeutic peptides, including antibacterial, anticancer, antifungal, and antiviral peptides (AntiBP3[p37] AntiFP[p38] AVPpred[p39] and AntiCP2[p40]). Similarly, *in silico* methods have been developed for the prediction of the half-life, toxicity, and structure of peptides (PlifePred[p41] ToxinPred2[p42] PepstrMod[p43] and HemoPI[p44]). Despite tremendous efforts in past two decades, there is still a huge gap between the number of known bioactive peptides and the number of peptides that have been approved for therapeutic use. Thus, there is a need to develop methods that can assist researchers to predict or design therapeutic peptides. Our database, THPdb2, is a tool that provides the information required for the development of prediction methods.

One of the primary features of our database is its comprehensive coverage of 354 mAb that are currently utilized in clinical practice. Most of these mAbs are used to treat cancer patients. Out of these 354 mAb drugs, only 129 mAbs have a publicly available amino acid sequence, and no sequence information is available for remaining 225 mAbs. A number of mAb drugs have failed in clinical trials. Thus, it is important to study properties of mAb drugs that are responsible for their failure in clinical trials. The mAbs in our database can be used to develop *in silico* models for identifying therapeutically viable and non-viable mAbs. In addition, our database contains 455 proteins other than mAbs such as biosimilar proteins, and the amino acid sequence is available for 214 of these proteins. The information on therapeutic proteins in our database will allow bioinformaticians to design protein drugs that can clear the different phases of clinical trials.

It is essential to acknowledge some of the limitations of our database, such as the unavailability of protein sequences for all entries and the absence of protein-based drugs that have been approved by international agencies other than the FDA. The THPdb2 data repository is manually curated and thoroughly verified to reduce the risk of error, but claiming complete accuracy would be unfair due to the possibility of human errors. Despite these limitations, our database remains a valuable resource for researchers in the field of therapeutic protein development.

With the increasing number of FDA-approved therapeutic peptide and protein candidates, the data in this field will also expand. We aim to update THPdb2 regularly as more FDA-approved therapeutic peptides and proteins become available in public repositories.

Conflict of interest

The authors declare no competing financial or non-financial interests.

Author contributions

SJ and SG collected and processed the data. SJ and SP created the back-end and front-end user interface of the web server. SJ, SG, SP and GPSR penned the manuscript. GPSR conceived and coordinated the project, and gave overall supervision to the project. All authors read and approved the final manuscript.

CRediT authorship contribution statement

Shipra Jain: Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Srijanee Gupta:** Writing – review & editing, Methodology, Data curation. **Sumeet Patiyal:** Writing – review & editing, Methodology. **Gajendra P.S. Raghava:** Validation, Supervision, Resources, Project administration, Conceptualization.

Data availability

Data will be made available on request.

Acknowledgement

SJ, SG and SP are thankful to the Department of Computational Biology, IIIT-Delhi for the provision of infrastructure and facilities.

References

1. Rossino G et al. Peptides as therapeutic agents: challenges and opportunities in the green transition era. *Molecules*. 2023;28:7165. https://doi.org/10.3390/molecules28207165.
2. Usmani SS, Kumar R, Bhalla S, Kumar V, Raghava GPS. In silico tools and databases for designing peptide-based vaccine and drugs. *Adv Protein Chem Struct Biol*. 2018;112:221–263. https://doi.org/10.1016/BS.APCSB.2018.01.006.
3. Sato AK, Viswanathan M, Kent RB, Wood CR. Therapeutic peptides: technological advances driving peptides into development. *Curr Opin Biotechnol*. 2006;17:638–642. https://doi.org/10.1016/J.COPBIO.2006.10.002.
4. Woodnutt G, Violand B, North M. Advances in protein therapeutics. *Curr Opin Drug Discov Dev*. 2008;11:754–761.

5. BioSpace. *Protein therapeutics market size to surpass USD 566.82 billion by 2030*. https://www.biospace.com/article/protein-therapeutics-market-size-to-surpass-usd-566-82-billion-by-2030/ [Published July 7, 2022. Accessed January 8, 2024].
6. Wang L et al. Therapeutic peptides: current applications and future directions. *Signal Transduct Target Ther*. 2022;7:48. https://doi.org/10.1038/S41392-022-00904-4.
7. Sharma N, Naorem LD, Gupta S, Raghava GPS. Computational resources in healthcare. *Wiley Interdiscip Rev Data Min Knowl Discov*. 2022;12:e1437.
8. Basith S, Manavalan B, Hwan Shin T, Lee G. Machine intelligence in peptide therapeutics: a next-generation tool for rapid disease screening. *Med Res Rev*. 2020;40:1276–1314. https://doi.org/10.1002/MED.21658.
9. Tyagi A et al. CancerPPD: a database of anticancer peptides and proteins. *Nucleic Acids Res*. 2015;43:D837–D843. https://doi.org/10.1093/NAR/GKU892.
10. Gawde U et al. CAMPR4: a database of natural and synthetic antimicrobial peptides. *Nucleic Acids Res*. 2023;51:D377–D383. https://doi.org/10.1093/NAR/GKAC933.
11. Wang G, Li X, Wang Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res*. 2016;44:D1087–D1093. https://doi.org/10.1093/NAR/GKV1278.
12. Usmani SS, Kumar R, Kumar V, Singh S, Raghava GPS. AntiTbPdb: a knowledgebase of anti-tubercular peptides. *Database (Oxford)*. 2018;2018, bay025. https://doi.org/10.1093/database/bay025.
13. Kumar R et al. AHTPDB: a comprehensive platform for analysis and presentation of antihypertensive peptides. *Nucleic Acids Res*. 2015;43:D956–D962. https://doi.org/10.1093/nar/gku1141.
14. Iwaniak A, Minkiewicz P, Darewicz M, Sieniawski K, Starowicz P. BIOPEP database of sensory peptides and amino acids. *Food Res Int*. 2016;85:155–161. https://doi.org/10.1016/J.FOODRES.2016.04.031.
15. Agrawal P et al. CPPsite 2.0: a repository of experimentally validated cell-penetrating peptides. *Nucleic Acids Res*. 2016;44:D1098–D1103. https://doi.org/10.1093/nar/gkv1266.
16. Kapoor P, Singh H, Gautam A, Chaudhary K, Kumar R, Raghava GPS. TumorHoPe: a database of tumor homing peptides. *PLoS One*. 2012;7:e35187.
17. Kumar V, Patiyal S, Kumar R, Sahai S, Kaur D, Lathwal A, Raghava GPS. B3Pdb: an archive of blood-brain barrier-penetrating peptides. *Brain Struct Funct*. 2021;226:2489–2495. https://doi.org/10.1007/S00429-021-02341-5.
18. Saha S, Bhasin M, Raghava GPS. Bcipep: a database of B-cell epitopes. *BMC Genomics*. 2005;6:79. https://doi.org/10.1186/1471-2164-6-79.
19. Lata S, Bhasin M, Raghava GP. MHCBN 4.0: a database of MHC/TAP binding peptides and T-cell epitopes. *BMC Res Notes*. 2009;2:61. https://doi.org/10.1186/1756-0500-2-61.
20. Vita R et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res*. 2019;47:D339–D343. https://doi.org/10.1093/nar/gky1006.
21. Gautam A et al. Hemolytik: a database of experimentally determined hemolytic and non-hemolytic peptides. *Nucleic Acids Res*. 2014;42:D444–D449. https://doi.org/10.1093/NAR/GKT1008.
22. van Ree R et al. The COMPARE Database: a public resource for allergen identification, adapted for continuous improvement. *Front Allergy*. 2021;2, 700533. https://doi.org/10.3389/falgy.2021.700533.
23. Mathur D et al. PEPlife: a repository of the half-life of peptides. *Sci Rep*. 2016;6:36617. https://doi.org/10.1038/srep36617.
24. Lefranc MP. IMGT, the International ImMunoGeneTics Information System for Immunoinformatics: methods for querying IMGT databases, tools, and web resources in the context of immunoinformatics. *Mol Biotechnol*. 2008;40:101–111. https://doi.org/10.1007/S12033-008-9062-7.
25. Manso T et al. IMGT® databases, related tools and web resources through three main axes of research and development. *Nucleic Acids Res*. 2022;50:D1262–D1272. https://doi.org/10.1093/nar/gkab1136.
26. Manso T, Kushwaha A, Abdollahi N, Duroux P, Giudicelli V, Kossida S. Mechanisms of action of monoclonal antibodies in oncology integrated in IMGT/mAb-DB. *Front Immunol*. 2023;14:1129323. https://doi.org/10.3389/fimmu.2023.1129323.
27. Usmani SS et al. THPdb: database of FDA-approved peptide and protein therapeutics. *PLoS One*. 2017;12, e0181748. https://doi.org/10.1371/journal.pone.0181748.
28. Wishart DS et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res*. 2006;34:D668–D672. https://doi.org/10.1093/nar/gkj067.
29. Kim S et al. PubChem 2023 update. *Nucleic Acids Res*. 2023;51:D1373–D1380. https://doi.org/10.1093/nar/gkac956.
30. Bateman A et al. UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res*. 2023;51:D523–D531. https://doi.org/10.1093/nar/gkac1052.
31. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
32. Smith TF, Waterman MS. Identification of common molecular subsequences. *J Mol Biol*. 1981;147:195–197. https://doi.org/10.1016/0022-2836(81)90087-5.
33. Fosgerau K, Hoffmann T. Peptide therapeutics: current status and future directions. *Drug Discov Today*. 2015;20:122–128. https://doi.org/10.1016/j.drudis.2014.10.003.
34. Mathieu C, Gillard P, Benhalima K. Insulin analogues in type 1 diabetes mellitus: getting better all the time. *Nat Rev Endocrinol*. 2017;13:385–399. https://doi.org/10.1038/NRENDO.2017.39.
35. Zaykov AN, Mayer JP, Dimarchi RD. Pursuit of a perfect insulin. *Nat Rev Drug Discov*. 2016;15:425–439. https://doi.org/10.1038/NRD.2015.36.
36. Sharma P, Joshi RV, Pritchard R, Xu K, Eicher MA. Therapeutic antibodies in medicine. *Molecules*. 2023;28:6438. https://doi.org/10.3390/molecules28186438.
37. Bajiya N, Choudhury S, Dhall A, Raghava GPS. AntiBP3: a method for predicting antibacterial peptides against Gram-positive/negative/variable bacteria. *Antibiotics (Basel)*. 2024;13:168. https://doi.org/10.3390/antibiotics13020168.
38. Agrawal P, Bhalla S, Chaudhary K, Kumar R, Sharma M, Raghava GPS. In silico approach for prediction of antifungal peptides. *Front Microbiol*. 2018;9:323. https://doi.org/10.3389/fmicb.2018.00323.
39. Thakur N, Qureshi A, Kumar M. AVPpred: collection and prediction of highly effective antiviral peptides. *Nucleic Acids Res*. 2012;40:W199–W204. https://doi.org/10.1093/nar/gks450.
40. Agrawal P, Bhagat D, Mahalwal M, Sharma N, Raghava GPS. AntiCP 2.0: an updated model for predicting anticancer peptides. *Brief Bioinform*. 2021;22, bbaa153. https://doi.org/10.1093/bib/bbaa153.
41. Mathur D, Singh S, Mehta A, Agrawal P, Raghava GPS. In silico approaches for predicting the half-life of natural and modified peptides in blood. *PLoS One*. 2018;13, e0196829. https://doi.org/10.1371/journal.pone.0196829.
42. Sharma N, Naorem LD, Jain S, Raghava GPS. ToxinPred2: an improved method for predicting toxicity of proteins. *Brief Bioinform*. 2022;23, bbac174. https://doi.org/10.1093/bib/bbac174.
43. Singh S et al. PEPstrMOD: structure prediction of peptides containing natural, non-natural and modified residues. *Biol Direct*. 2015;10:73. https://doi.org/10.1186/s13062-015-0103-4.
44. Kumar V, Kumar R, Agrawal P, Patiyal S, Raghava GPS. A method for predicting hemolytic potency of chemically modified peptides from its structure. *Front Pharmacol*. 2020;11:54. https://doi.org/10.3389/fphar.2020.00054.