

2015

Designing AMPed: The Practical Antimicrobial Peptide Editable Database

Tripti Garg
University of Rhode Island

George Konstantinidis
University of Rhode Island

See next page for additional authors

Follow this and additional works at: https://digitalcommons.uri.edu/cmb_facpubs

Terms of Use
All rights reserved under copyright.

Citation/Publisher Attribution

Garg, T., Konstantinidis, G., Peckham, J., Monteiro, A., LaCroix, R., & Martin, L. M. (2015). Designing AMPed: The Practical Antimicrobial Peptide Editable Database. In V. Srivastava, A. Yudin, & M. Lebl (Eds.). *Proceedings of the 24th American Peptide Symposium* (pp. 105-108). Orlando, FL. Retrieved from <http://dx.doi.org/10.17952/24APS.2015.105>
Available at: <http://dx.doi.org/10.17952/24APS.2015.105>

This Article is brought to you for free and open access by the Cell and Molecular Biology at DigitalCommons@URI. It has been accepted for inclusion in Cell and Molecular Biology Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.

Authors

Tripti Garg, George Konstantinidis, Joan Peckham, Admir Monteiro, Roxanne LaCroix, and Lenore M. Martin

Designing AMPed: The Practical Antimicrobial Peptide Editable Database

**Tripti Garg^{1,2}, George Konstantinidis², Joan Peckham², Admir Monteiro³,
Roxanne LaCroix¹, and Lenore M. Martin¹**

¹Dept. of Cell & Molecular Biology; ²Dept. of Computer Science & Statistics; ³Electrical, Computer, & Biomedical Engineering, The University of Rhode Island, Kingston, RI, 02881, USA

Introduction

Peptide data are stored in various online repositories. Researchers typically need to sift through numerous databases containing large quantities of unwanted information before locating the exact data they need. Routine access to vital data is a very complex and daunting challenge. Individual researchers, experimenting with innovative search and data filtering strategies, have built up their own local databases, adding to an ever-increasing volume of highly dispersed, partially overlapping, peptide datasets. None of the existing online databases uniformly annotates their data, making correlating entries for identical proteins or peptides from one dataset to another a staggering task. Bioinformaticians have developed software to simplify cross-database searching, but the problem of locating useful antimicrobial peptide research-related entries buried inside general databases remains. Presently, AMPed consists of data obtained by bulk downloads of proteins less than 100 residues from four large online portals: UniProt (www.uniprot.org), NCBI (www.ncbi.nlm.nih.gov), EBI (www.ebi.ac.uk), and KEGG (www.genome.jp).

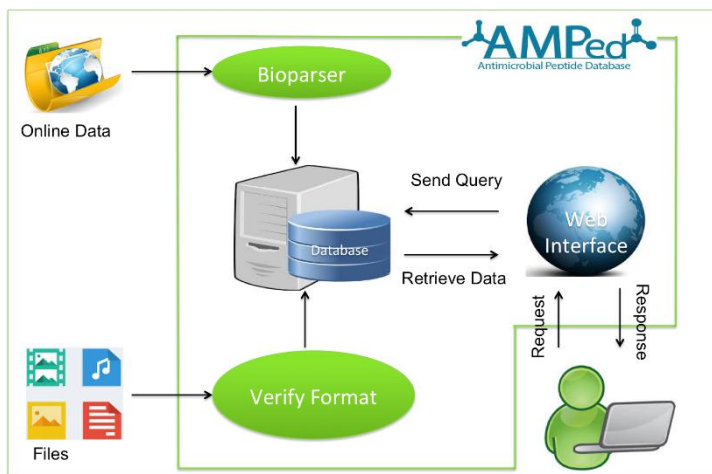


Fig. 1. The overall view of the AMPed system.

AMPed is under development at URI to support our ongoing research program on antimicrobial peptides. This work will enhance antimicrobial data access, facilitate long-distance collaborations, improve the search speed, and eliminate the need to manually correlate information gathered from multiple data sources, thereby greatly reducing the time researchers have to spend to locate and evaluate the desired data sets.

The integration of Bioparser, the AMPed database, and its web-based Interface is defined as the AMPed system (Figure 1). Bioparser, built in GO language, parses the bulk data downloaded from online repositories, converts it into a harmonious format and builds some connections. The AMPed database, developed in MySQL, is a highly normalized relational database that stores peptide data. The secure web-based interface, built using PHP & HTML5, provides easy access to the AMPed database

in the cloud. The new intuitive web interface is designed to help researchers easily formulate combinations of frequent database queries to meet their specific needs.

Navigation: A user-friendly interface for AMPed works equally well for both novice and expert users. The screens are organized logically to help users perform tasks efficiently. The navigation scheme for the AMPed interface is depicted in Figure 2.

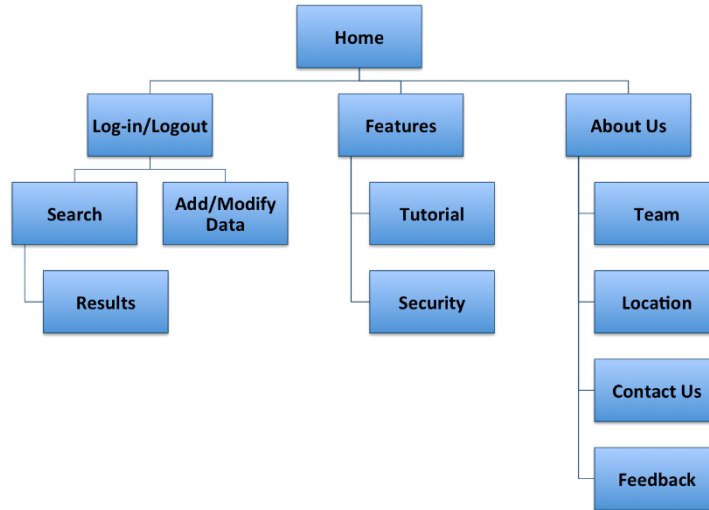


Fig. 2. Displays navigation scheme of webpages of the AMPed web interface.

The latest focus of the AMPed development project, is to build an easy, secure and intuitive user interface suitable for novice through expert researchers to efficiently search, display, manipulate, and eventually upload their own data into AMPed. The user interface project incorporates insights into various aspects of use of the current data set, and adds on and evaluates the utility of diverse visual displays of data.

Interface Design: There have been several empirical studies that identified basic psychological factors that should be considered when designing a Graphical User Interface (GUI). To design the AMPED GUI, we considered the following three primary contributing human factors:

- **Physical limits of visual acuity** - Visual acuity is the ability of the eye to resolve detail. The retina of the human eye can only focus on a small portion of a computer screen. The AMPed GUI will limit the size of icons, menus, dialog boxes etc. to ensure they fit into the limited area the human eye can take in at any one time. The design will also group related material to maintain user focus on one section of the screen. This will ensure that the user does not have to repeatedly move their eyes across the screen, causing eye fatigue due to unnecessary movements.
- **Limits of absolute memory** - Once the user has a desired fixation point, there is a limit to the amount of information that the person can store and process at one viewing. A GUI design “rule of thumb” is that the range of options or choices should never exceed five or six. AMPed GUI will chunk the information presented to the user (e.g. search results) into logical groups and will ensure not to crowd screen with data. It will also optimize the number of menu options (e.g. number of search criteria) on the page to develop clean and simple intuitive designs.
- **Gestalt Principle** - The Gestalt Principle states that people use a top-down approach to organizing data. AMPed GUI will leverage the Gestalt principle of ‘continuation’ by organizing the data in a top-down approach. It will leverage the principle of ‘similarity’ via using a template-driven approach to help users easily find like items, for example home link, log-in, log-out and search functions, at a consistent place across all screens. It will use the principle of ‘proximity’ to place similar search results close together so they are perceived as a logical group.

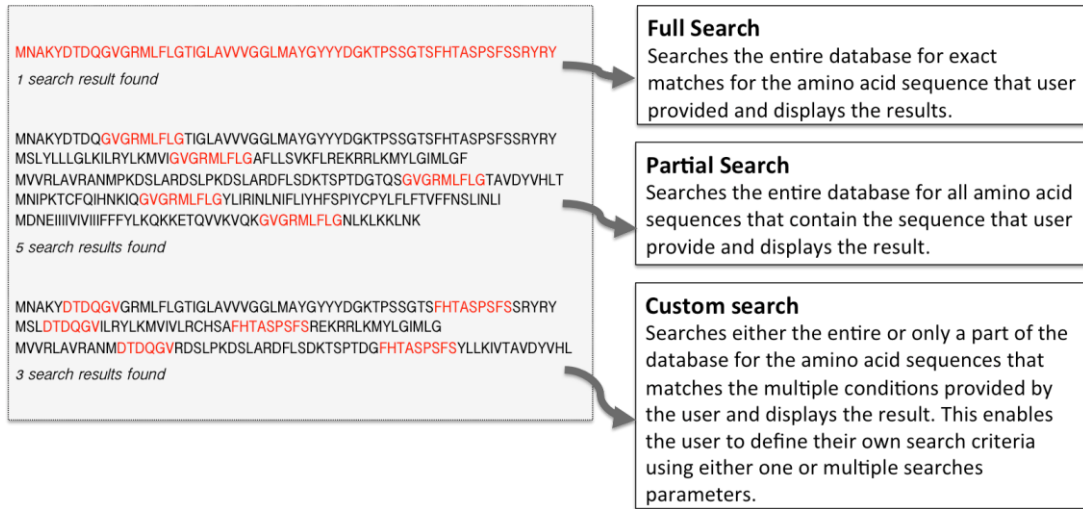


Fig. 3. A view of amino acid sequence search results.

Search Types: The full to custom search option allows researchers to easily parse through huge amounts of data quickly, efficiently and accurately. The example above (Figure 3) highlights how a full, partial or custom string search using an amino acid sequence provides different results when searched via AMPed.

Search Queries: The user interface provides the ability for a user to choose search strategies ranging from simple queries to complex combinations. The search function within AMPed consists of multiple pre-written SQL queries that are optimized for speed and accuracy. The results are sorted and viewed directly in the webpage or reports are offered that may be downloaded in multiple formats.

Data Security: The AMPed interface is designed to be highly secure and to meet the needs of different users by incorporating the following security mechanisms:

- Provides secure access to ensure data integrity and proper user authorization.
- Has CAPTCHA to protect from DOS (denial of service) attacks by robotic scripts.
- Provides tiered access levels for users: professors, students, experts & creators.
- Maintains access and audit logs for system managers to observe user trends

AMPed Website: An open source, user-friendly and well-secured website meant to provide access to the AMPed database by novice and expert users. The idea of “AMPed” (Anti-Microbial Peptide Editable Database) is to create a targeted annotated database consisting of uniquely focused information about antimicrobial peptides with the primary purpose of data storage uniformity and coherence. AMPed will be an open source public database with a simple and secure web interface that peptide researchers can use to locate and download sequences, structures, and biological information relevant to antimicrobial research, linked to the original data sources. AMPed is powered by a software tool called “Bioparser” that automatically extracts, sorts, and allows users to annotate data derived from online repositories.

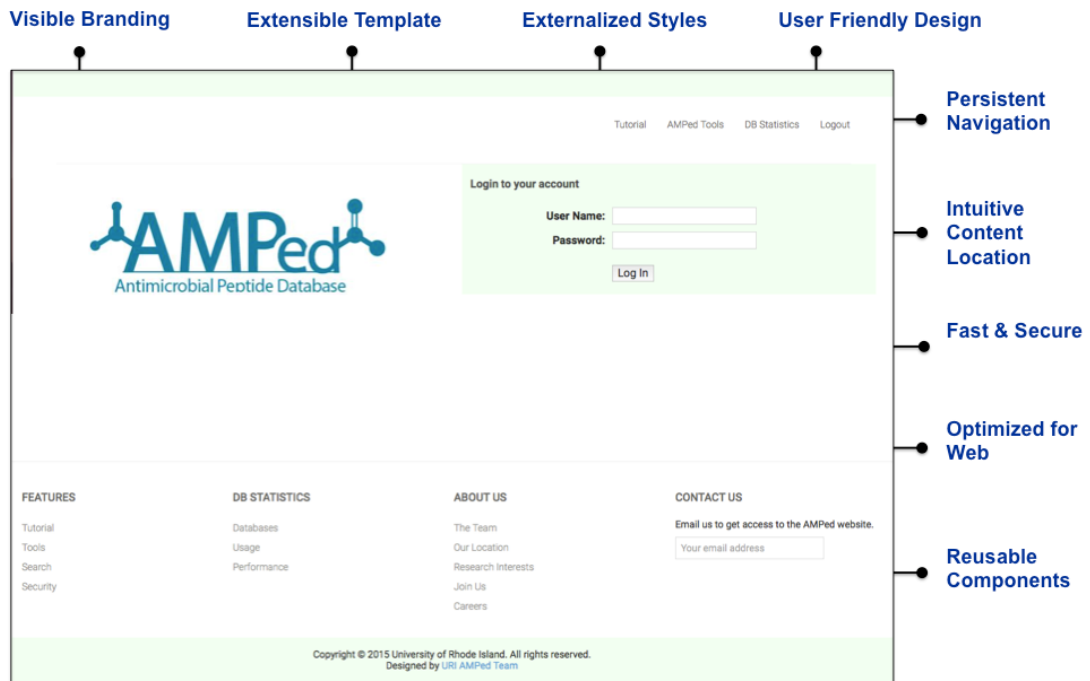


Fig. 4. The Home webpage of the AMPed web based user interface.

Acknowledgments

George Konstantinidis for designing Bioparser, & population of the AMPed database. BCH-CSC522 students for helping us take the first steps in this project. Professors Lenore M. Martin, Jean-Yves Hervé & Joan Peckham for valuable guidance & advice.

References

1. 2012, www.bbcm.univ.trieste.it/~tossi/pag1.htm, AMSDb, A Tossi, Trieste.
2. 2015, www.peptideatlas.org, Peptide Atlas, E Deutsch, Seattle.
3. 2012, aps.unmc.edu/AP/main.php, Antimicrobial Peptide Database, G Wang, Nebraska
4. 2012, 159.226.149.45/other1/kizapd/, Antimicrobial Peptide Database, Kunming Institute of Zoology, China
5. 2015, *Nucleic Acids Research*, Vol 43, Database Issue
6. 2014, George Konstantinidis; Thesis
7. 1998, E Tufte; *Visual Explanation: Images & Quantities, Evidence & Narrative*, CT
8. *International Journal of Man-Machine Studies*, Benbasat, I& P Todd; An Experimental Investigation of Interface Design Alternatives
9. 2010, Roman Pichler; *Agile Product Management with Scrum: Creating Products that Customers Love*, Addison-Wesley Professional, Boston