D590–D592 *Nucleic Acids Research, 2004, Vol. 32, Database issue*
DOI: 10.1093/nar/gkh025

APD: the Antimicrobial Peptide Database

Zhe Wang and Guangshun Wang*

Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, 986805 Nebraska Medical Center, Omaha, NE 68198-6805, USA

Received August 11, 2003; Revised August 20, 2003; Accepted September 3, 2003

ABSTRACT

An antimicrobial peptide database (APD) has been established based on an extensive literature search. It contains detailed information for 525 peptides (498 antibacterial, 155 antifungal, 28 antiviral and 18 antitumor). APD provides interactive interfaces for peptide query, prediction and design. It also provides statistical data for a select group of or all the peptides in the database. Peptide information can be searched using keywords such as peptide name, ID, length, net charge, hydrophobic percentage, key residue, unique sequence motif, structure and activity. APD is a useful tool for studying the structure–function relationship of antimicrobial peptides. The database can be accessed via a web-based browser at the URL: http://aps.unmc.edu/AP/main.html.

INTRODUCTION

Antimicrobial peptides have been identified in various species ranging from bacteria, frogs to mammals, including humans. They form the first line of host defense against pathogenic infections and are a key component of the ancient innate immune system. Most antimicrobial peptides possess 6–50 amino acid residues with net positive charges (1–3). It is generally accepted that these cationic peptides selectively interact with anionic bacterial membranes (4,5) although different mechanisms may be used by different peptides under different conditions for killing (6). Currently, there is great interest in antimicrobial peptides because these so-called 'nature's antibiotics' are promising to overcome the growing problems of antibiotic resistance (1–3).

Until now, more than 500 antimicrobial peptides have been reported (1,2). Although some structural and sequence information of these peptides can be retrieved from PDB (7) and Swiss-Prot (8), the majority of the other antimicrobial data are scattered in the literature. This is not convenient for a more comprehensive structural and functional study of these peptides. For this reason, we have created a database to store and analyze these interesting peptides.

DATABASE CONSTRUCTION

APD was built on the Red Hat Linux operating system using the freeware Apache web server, PHP script language and MySQL relational database system. Antimicrobial peptides were collected from the literature by PubMed search using keywords such as 'antimicrobial peptide', 'antibacterial peptide', 'antifungal peptide', 'anticancer peptide' or 'antitumor peptide'. The peptides collected in this version of APD are mainly from natural sources. For the purpose of database bookkeeping, a unique five-digit identification number (ID) starting with AP was assigned to each peptide. Each entry was checked in PDB or Swiss-Prot. If the peptide exists in the established databases, a web link was created in APD for that entry to facilitate consultation of the original databases. In addition, each entry also contains peptide name, sequence, length, hydrophobic percentage, net charge, structure (such as α-helix or β-strand), physical method for structural determination (e.g. NMR spectroscopy or X-ray diffraction), biological activity (antibacterial, antiviral, antifungal or antitumor), critical residue for activity, and reference (including authors, article title, journal, page, volume and year).

DATABASE DESCRIPTION

The database main page contains the following interfaces: About (introduction to APD), Database (the query interface), Prediction, Peptide design, Statistical data, Useful links and Contact information. There are also links for antibacterial peptides, antiviral peptides, antifungal peptides and anticancer peptides. Each link contains a list of the peptides with a specific function.

The Query interface

Users can click on the Database icon and the search interface appears. APD can then be searched in multiple ways. As a dictionary, the user can find out whether a specific peptide is collected in APD by entering the peptide name or the N-terminal sequence of the peptide. When one or two sequence motifs are entered, the program will return all peptides containing those motifs. For example, there are three peptides containing the sequence LysLysLysLys, but no peptide contains the sequence ArgArgArgArg in the current version of APD. Users can also do a search by using PDB ID, Swiss-Prot ID (if there is one), structure, structural method, hydrophobic percentage, net charge, length and activity. The search results can be sorted by peptide length, net charge, ID, or by peptide location such as PDB, Swiss-Prot or the literature. Detailed information for each entry can be viewed by clicking on the peptide ID.

*To whom correspondence should be addressed. Tel: +1 402 559 4176; Fax: +1 402 559 4651; Email: gwang@unmc.edu

Table 1. Peptide structural statistics in APD (total 525 entries)

Peptide structure	Number	Percentage
α-Helical structure	77	14.67
β-Structure without disulfide bonds	4	0.76
Both α-helix and β-structure	13	2.47
Rich in specific residues	53	10.09
β-Structure with disulfide bonds	170	32.38
No known structure	208	39.62

The Prediction and Peptide design interfaces

The Prediction and Peptide design interfaces were generated as a database-based tool for designing novel peptides with required functions. In APD, hydrophobic residues are defined according to the Kyte–Doolittle scale (9). Trp is also counted as a hydrophobic residue because of its importance in lipid binding (10,11) and preference in the protein–membrane interface (12). The Prediction interface allows users to input a new peptide sequence. The program will carry out a residue analysis on the peptide. It also predicts whether the new peptide has the potential to be antimicrobial based on some known principles. In terms of structure, only some simple predictions can be made. For instance, when the hydrophobic residues appear every two to three residues in the peptide sequence, an amphipathic helix will be predicted. The sequence alignment function can be initiated in either the Prediction or the Peptide design interface. The system will perform alignments between the input sequence and those sequences in the database. Five peptide sequences most similar to the user's input will be displayed. The sequence alignment results are represented in pairs with the input sequence always shown underneath for comparison.

The Statistical interface

This interface provides statistical data on peptide sequence, function and structure. For sequence analysis, the average length, net charge and residue percentage of all peptides in the database are listed. For functional analysis, the numbers (percentages) of antimicrobial, antifungal, antiviral and anticancer peptides are given. For structural analysis, the number of peptides with a defined structural type will be shown. Users can also get sequence statistical data for any search results.

APD CURRENT HOLDINGS AND FINDINGS

The current version of APD holds 525 antimicrobial peptides. Among them, 498 have an effect on bacteria, 155 on fungi, 28 on viruses and 18 on cancer cells. Note that a specific peptide may have different functions. Thus, it can be counted twice or more. Based on the structure and sequence features, these peptides have been classified into six groups in APD. Table 1 lists the number of the peptides in each group. Clearly, the most abundant peptides are those with disulfide bonds or those without known structures. Indeed, only 54 antimicrobial peptides were found to have known 3D structures deposited in the PDB. In our database, however, we collected 68 peptides that were studied by NMR spectroscopy in membrane-mimetic environments or lipid bilayers, whereas only a few crystal structures were found, indicating that NMR is the major player in this game.

Table 2. Antimicrobial peptides with different functions have different amino acid residue ratios

%	All[a]	AB	AF	AV	AC	TO
Ile	6.82	7.03	5.42	5.28	9.00	9.31
Val	6.20	6.38	4.90	5.42	6.60	6.73
Leu	8.86	9.18	7.40	6.11	9.60	12.00
Phe	4.39	4.45	3.86	4.17	7.80	3.53
Cys	6.57	6.23	8.90	15.99	1.20	3.25
Met	0.89	0.90	0.97	0.55	0.90	0.67
Ala	8.21	8.40	8.09	6.39	7.50	11.16
Trp	1.41	1.36	1.82	2.08	2.40	1.62
Pho	*43.35*	*43.93*	*41.36*	*45.99*	*45.00*	*48.27*
Gly	11.10	10.86	11.11	10.01	8.40	10.77
Pro	4.75	4.77	3.95	3.75	2.10	3.19
Thr	3.46	3.42	3.51	3.75	2.40	3.19
Ser	5.03	4.98	5.44	4.03	4.80	5.05
Tyr	2.20	2.12	3.36	4.86	1.20	0.61
Gln	2.58	2.51	2.93	2.08	1.80	2.02
Asn	3.02	2.85	3.01	2.50	0.30	2.74
Glu	2.04	2.03	2.17	2.36	1.20	2.02
Asp	1.95	1.93	1.84	0.55	2.10	1.73
His	2.38	2.23	2.43	1.80	3.90	2.13
Lys	10.60	10.68	9.85	2.22	16.21	13.97
Arg	7.32	7.44	8.57	15.99	5.40	4.20

[a]All: all antimicrobial peptides stored in the database (525 peptides: average length 27.97, average net charge 4.56); AB: antibacterial peptides (498 peptides: average length 27.82, average net charge 4.56); AF: antifungal peptides (155 peptides: average length 29.72, average net charge 5.01); AV: antiviral peptides (28 peptides: average length 25.68, average net charge 4.39); AC: anticancer peptides (18 peptides: average length 18.50, average net charge 4.11); TO: peptides have a toxic effect on mammalian cells (64 peptides: average length 27.84, average net charge 4.61); Pho: all hydrophobic residues as defined in the text.

In APD, 97% of the peptides contain 50 residues or less with an average length of 28. The longest peptide (AP00404) collected has 84 residues and the shortest one (AP00027) merely six residues. The majority of peptides (96%) in the database have a net positive charge with the highest being +17 (AP00010). As a result, the average net charge of all peptides in APD is +4.6. (Note that the effect of chemical modification and pH on the peptide charge has not been programmed in this version.) Some antimicrobial peptides, however, have a net negative charge. The most negatively charged peptide contains a string of Asp residues with a net charge of –6 (AP00528). It requires zinc as a cofactor for activity (13). Another negatively charged peptide in APD is maximum H5, which does not require a cofactor (14). It contains both hydrophobic and three Asp residues but no cationic residues. This anionic antimicrobial peptide is promising to kill the human β-defensin-resistant Gram-positive bacterium *Staphylococcus aureus*, which escapes attacks from cationic peptides probably by incorporating positive charges on the membrane surface by adding Lys to lipids (15). In light of the ancient Chinese ying-yang philosophy, the anionic antimicrobial peptides, although rarely documented (13,14), appear to complement the cationic antimicrobial peptides, offering us a complete spectrum of antimicrobial peptides.

The average contents of hydrophobic residues in different functional groups of the peptides are similar, ranging from 41% to 49% (Table 2, italicized). Although the populations of antiviral and antitumor peptides are relatively small in the

database, it is interesting to note that antiviral peptides have the highest Cys content (16.0 versus average 6.6), causing 78% of the peptides in this class to have one or more disulfide bonds. Also, they have a lower ratio of Lys (2.2 versus average 10.6) but a higher ratio of Arg (16.0 versus average 7.3). In strong contrast, anticancer peptides have a lower Cys ratio but a relatively high content of Lys. Of note, antimicrobial peptides that showed a toxic effect on mammalian cells have slightly higher contents of Ile, Leu and Ala in the sequence, leading to the highest hydrophobic content (Table 2). These findings may be useful in improving natural peptide templates or designing novel peptides with minimal side effects on humans.

AVAILABILITY

APD is created, maintained and updated by the NMR Structural Biology Group at Eppley Institute, the University of Nebraska Medical Center. It can be accessed via the Internet at the URL: http://aps.unmc.edu/AP/main.html. Researchers in this field are invited to use the database, make suggestions and submit their peptides by contacting the authors. The database will be expanded and improved in our next release.

ACKNOWLEDGEMENTS

We thank Zhengxin Chen, Hesham H. Ali and J. Philip Craiger (University of Nebraska at Omaha) for advice on database design and programming, Atul Rayamajhi and Joe Ziskovsky (UNMC) for assistance in testing and activating the online version of the database, and Wally Murphy and Bill Goodrich (UNMC) for computer support. We are also grateful to Kristi Berger (UNMC) for text editing. G.W. acknowledges financial support from the Eppley Institute, UNMC, for this research.

REFERENCES

1. Hancock,R.E.W. and Patrzykat,A. (2002) Clinical development of cationic antimicrobial peptides: from natural to novel antibiotics. *Curr. Drug Targets Infect. Disord.*, **2**, 79–83.
2. Scott,M.G. and Hancock,R.E.W. (2000) Cationic antimicrobial peptides and their multifunctional role in the immune system. *Crit. Rev. Immunol.*, **20**, 407–431.
3. Bradshaw,J.P. (2003) Cationic antimicrobial peptides: issues for potential clinical use. *BioDrugs*, **17**, 233–240.
4. Marcotte,I., Wegener,K.L., Lam,Y.H., Chia,B.C., de Planque,M.R., Bowie,J.H., Auger,M. and Separovic,F. (2003) Interaction of antimicrobial peptides from Australian amphibians with lipid membranes. *Chem. Phys. Lipids*, **122**, 107–120.
5. Hancock,R.E.W. and Rozek,A. (2002) Role of membranes in the activities of antimicrobial cationic peptides. *FEMS Microbiol. Lett.*, **206**, 143–149.
6. Epand,R.M. and Vogel,H.J. (1999) Diversity of antimicrobial peptides and their mechanisms of action. *Biochim. Biophys. Acta*, **1462**, 11–28.
7. Berman,H.M. Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
8. Bairoch,A. and Apweiler,R. (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.*, **28**, 45–48.
9. Kyte,J. and Doolittle,R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.*, **157**, 105–132.
10. Wang,G., Pierens,G.K., Treleaven,W.D., Sparrow,J.T. and Cushley,R.J. (1996) Conformations of human apolipoprotein E(263–268) and E(267–289) in aqueous solutions of sodium dodecyl sulfate by CD and ^{1}H NMR. *Biochemistry*, **35**, 10358–10366.
11. Wang,G. (2002) How the lipid-free structure of the N-terminal truncated human apoA-I converts to the lipid-bound form: new insights from NMR and X-ray structural comparison. *FEBS Lett.*, **529**, 157–161.
12. Wimley,W.C. and White,S.H. (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature Struct. Biol.*, **3**, 842–848.
13. Brogden,K.A., De Lucca,A.J., Bland,J. and Elliott,S. (1996) Isolation of an ovine pulmonary surfactant-associated anionic peptide bactericidal for *Pasteurella haemolytica*. *Proc. Natl Acad. Sci. USA*, **93**, 412–416.
14. Lai,R., Liu,H., Lee,W.H. and Zhang,Y. (2002) An anionic antimicrobial peptide from toad *Bombina maxima*. *Biochem. Biophys. Res. Commun.*, **295**, 796–799.
15. Peschel,A., Jack,R.W., Otto,M., Collins,L.V., Staubitz,P., Nicholson,G., Kalbacher,H., Nieuwenhuizen,W.F., Jung,G., Tarkowski,A., van Kessel,K.P.M. and van Strijp,J.A.G. (2001) *Staphylococcus aureus* resistance to human defensins and evasion of neutrophil killing via the novel virulence factor MprF is based on modification of membrane-lipids with L-lysine. *J. Exp. Med.*, **193**, 1067–1076.