

DADP: the database of anuran defense peptides

Mario Novković¹, Juraj Simunić¹, Viktor Bojović², Alessandro Tossi³ and Davor Juretić^{1,*}

¹Department of Physics, Faculty of Science, University of Split, 21000 Split, ²Centre for Informatics and Computing, Ruđer Bošković Institute, 10000 Zagreb, Croatia and ³Department of Life Sciences, University of Trieste, 34127 Trieste, Italy

Associate Editor: Burkhard Rost

ABSTRACT

Summary: Anuran tissues, and especially skin, are a rich source of bioactive peptides and their precursors. We here present a manually curated database of antimicrobial and other defense peptides with a total of 2571 entries, most of them in the precursor form with demarcated signal peptide (SP), acidic proregion(s) and bioactive moiety(s) corresponding to 1923 non-identical bioactive sequences. Search functions on the corresponding web server facilitate the extraction of six distinct SP classes. The more conserved of these can be used for searching cDNA and UniProtKB databases for potential bioactive peptides, for creating PROSITE search patterns, and for phylogenetic analysis.

Availability: DADP is accessible at <http://split4.pmfst.hr/dadp/>

Contact: juretic@pmfst.hr

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on December 6, 2011; revised on March 2, 2012; accepted on March 20, 2012

1 INTRODUCTION

With the growing problem of pathogens developing resistance to conventional antibiotics, there is a great interest in antimicrobial peptides (AMPs) as substitute therapeutics (Hancock and Sahl, 2006). There are several existing databases on many types of AMPs: APD (Wang *et al.*, 2009), CAMP (Thomas *et al.*, 2010), PhytAMP (Hammami *et al.*, 2009), Defensins Knowledgebase (Seebah *et al.*, 2007), AMSDB (Tossi and Sandri, 2002), BACTIBASE (Hammami *et al.*, 2007), AMPPer (Fjell *et al.* 2007) and YADAMP (Piotto *et al.* 2012). However, these databases tend to focus on the antimicrobial sequences, and the immature precursors from which the active AMPs are cleaved are not always present. DADP is a versatile database for anuran defense peptides that retains all this important information. Some anuran species contain >100 AMPs in their tissues, and especially their skin, as well as other bioactive peptides (Li *et al.*, 2007, Yang *et al.*, 2012), but many are endangered, while whole genomes have been decoded for only 2 out of almost 6000 species (Roelants *et al.*, 2007; Calboli *et al.* 2011).

2 DATA COLLECTION

The following query was used in UniProtKB (<http://www.uniprot.org/>, Apweiler *et al.*, 2011) release 2012_02 (February 22, 2012): (keyword: “Amphibian defense peptide [KW-0878]” OR

keyword: “Antimicrobial [KW-0929]” OR keyword: “Antibiotic [KW-0044]”) AND taxonomy: “Amphibia [8292]” and 2492 results were obtained. All sequences denoted as fragments were omitted. For any region that was not annotated in UniProtKB, the sequence was manually curated and divided into specific regions (signal, acidic and bioactive) based on homology with annotated protein precursors coupled with SignalP prediction (Bendtsen *et al.*, 2004). Several protein sequences not in accord with cited literature were omitted.

The final version of the database contains 1766 full precursor sequences and 805 bioactive peptides with unknown precursor, encompassing a total of 167 anuran species from 12 families (Supplementary Tables S1 and S2). Biogeography data were collected using The IUCN Red List of Threatened Species (www.iucnredlist.org) and Amphibian Species of the World 5.5, American Museum of Natural History (<http://research.amnh.org/vz/herpetology/amphibia/index.php>).

Classification by biological function includes over 70 different activities, with AMPs being in the majority (Supplementary Table S3). Missing activities field denotes unknown function. Phylogenetic analysis and classification in six classes, based on signal peptide (SP) similarity (Supplementary Fig. S1), was constructed using *MEGA* software version 5 (Tamura *et al.*, 2007). Activity and cytotoxicity data were extracted from published papers where the minimal inhibitory concentration (MIC, μ M) based on microdilution assays was reported for *Escherichia coli* and/or *Staphylococcus aureus*. HC₅₀ refers to peptide concentration, also micromolar, required for 50% haemolysis of red blood cells. MIC data are available for 921 peptide sequences.

3 IMPLEMENTATION

DADP is hosted on a 64 bit Linux server using Apache (<http://www.apache.org/>) and PHP (<http://www.php.net/>). All entries are stored in a MySQL (<http://www.mysql.com/>) database. The application was written in PHP, using the jQuery JavaScript Library (<http://jquery.com/>) with the DataTables (<http://datatables.net/>) plug-in for the tables' implementation, and CSS for web design.

4 WEB INTERFACE

The DADP web interface offers browsing capability for all peptide entries, for SPs and for SP classes in particular. An advanced search option is available for more specific user requests. Each DADP ID entry contains detailed information about peptide names and sequences, tripartite structure, species, SP class, biogeography,

*To whom correspondence should be addressed.

biological function, peptide activity and links to UniProtKB. Search results are presented as tables.

5 DISCUSSION

Bioactive peptides from amphibian species with function, often related to host defense, are a fast growing group in the UniProtKB database and this attempt to collect and order them all can mostly serve to emphasize the incompleteness of our knowledge. For instance, out of 167 species from the DADP, only 28 can be found among 2200 threatened and near threatened anuran species from the IUCN Red List. The DADP focuses on precursors and their conserved SPs and has allowed the individuation of six distinct SP classes (see Supplementary Material and Supplementary Tables S4 and S5). Consensus sequences for the different SP classes are a good foundation for searching protein or DNA databases that might carry precursors with novel AMP sequences, where a direct homology search based on the active region only is unsuccessful (Juretić *et al.*, 2011; Tessera *et al.*, 2012). The most numerous class of SPs (present in 85% of precursors) has twin lysines at positions 5 and 6, and is denoted as Class-1 in DADP. It is found in Neobatrachia families Ranidae, Hylidae, Hyperoliidae and Rhacophoridae. SP classes 1–6 are associated not only with AMPs but also with mature bioactive peptides having other functions.

Precursor structure is informative about origin, function and evolution of mature peptides collected in the DADP, but there are numerous peptides present only as the mature bioactive sequence, where the precursor structure is not reported. The activity data, shown as MIC and HC₅₀ values, are available for over 900 sequences. These should, however, be considered critically as both are quite sensitive to assay conditions, which vary from lab to lab [e.g. medium, salt concentrations and the size of initial inoculum, which should be at least 10⁵ cfu/ml for MIC (Resende *et al.*, 2008) and at most 1% erythrocytes for HC₅₀ determinations, limits often either not observed or not defined in published papers].

The conservation of SPs in different suborders and families of anurans, whose relatedness follows continental tectonics and corresponding changes in biogeography, is quite impressive (Supplementary Fig. S1 and the DADP web server). This is one of those rare examples when maintenance of the export SPs, which are dispensed with after secretion, is evolutionarily more conserved than the mature exported polypeptide structure. It may be beneficial for the host to have accelerated co-evolution of the mature defense peptide sequences, to keep up with the constantly shifting microbial biota, allowing to maintain the host-pathogen balance (Peschel and Sahl, 2006). The number of identical SPs associated with different mature peptides found in different species can be quite large (e.g. 164 identical Class-1 SP sequences encompassing 36 species and associated with 141 different mature peptides).

It is also possible to find identical mature peptides associated with SPs that are different in several positions. An example is the antimicrobial granulosisin-D2 and palustrin-2ISa, respectively, belonging to different species (*Amolops granulatus* from China and *Rana ishikawae* from Japan). Another is the bradykinin sequence RPPGFSPFR, which is associated with different SP sequences in 15 different species. These bioactive skin peptides from frogs and toads may be effective repellents against macroscopic predators such as

fish (Chen *et al.*, 2011) and snakes. These two cases are possible examples of convergent evolution of the active region to provide for similar functions.

The DADP can greatly facilitate using the HMMER tool (Finn *et al.*, 2011) to search UniProtKB for amphibian AMPs. In this case the sequence length limitations of the tool require the combined use of the signal sequence and acidic propiece as the query.

The DADP can be regularly updated using both keywords and similarity search.

ACKNOWLEDGEMENTS

Funding: Croatian Ministry of Science, Education and Sport (Grant No. 177-1770495-0476).

Conflict of Interest: none declared.

REFERENCES

- Apweiler, R. *et al.* (2011) Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res.*, **39**, 214–219.
- Bendtsen, J.D. *et al.* (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.*, **340**, 783–795.
- Calboli, F.C. *et al.* (2011) The need for jumpstarting amphibian genome projects. *Trends Ecol. Evol.*, **26**, 378–379.
- Chen, X. *et al.* (2011) A fish bradykinin (Arg⁰, Trp⁵, Leu⁸-bradykinin) from the defensive skin secretion of the European edible frog, *Pelophylax kl. esculentus*: structural characterization; molecular cloning of skin kininogen cDNA and pharmacological effects on mammalian smooth muscle. *Peptides*, **32**, 26–30.
- Finn, R.D. *et al.* (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.*, **39**, W29–W37.
- Fjell, C.D. *et al.* (2007) AMPper : a database and an automated discovery tool for antimicrobial peptides. *Bioinformatics*, **23**, 1148–1155.
- Hammami, R. *et al.* (2007) BACTIBASE: a new web-accessible database for bacteriocin characterization. *BMC Microbiol.*, **7**, 89.
- Hammami, R. *et al.* (2009) PhytAMP: a database dedicated to antimicrobial plant peptides. *Nucleic Acids Res.*, **37**, 963–968.
- Hancock, R.E. and Sahl, H.G. (2006) Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nat. Biotechnol.*, **24**, 1551–1557.
- Juretić, D. *et al.* (2011) Knowledge-based computational methods for identifying or designing novel, non-homologous antimicrobial peptides. *Eur. Biophys. J.*, **40**, 371–385.
- Li, J. *et al.* (2007) Anti-infection peptidomics of amphibian skin. *Mol. Cell Proteomics*, **6**, 882–894.
- Peschel, A. and Sahl, H.G. (2006) The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nat. Rev. Microbiol.*, **4**, 529–536.
- Piotto, S.P. *et al.* (2012) YADAMP: yet another database of antimicrobial peptides. *Int. J. Antimicrob. Agents*, **39**, 346–351.
- Resende, J.M. *et al.* (2008) Solution NMR structures of the antimicrobial peptides phylloseptin-1, -2, and -3 and biological activity: the role of charges and hydrogen bonding interactions in stabilizing helix conformations. *Peptides*, **29**, 1633–1644.
- Roelants, K. *et al.* (2007) Global patterns of diversification in the history of modern amphibians. *Proc. Natl Acad. Sci. USA*, **104**, 887–892.
- Seebah, S. *et al.* (2007) Defensins knowledgebase: a manually curated database and information source focused on the defensins family of antimicrobial peptides. *Nucleic Acids Res.*, **35**, 265–268.
- Tamura, K. *et al.* (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.*, **24**, 1596–1599.
- Tessera, V. *et al.* (2012) Identification of antimicrobial peptides from teleosts and anurans in EST databases using conserved signal sequences. *FESB J.*, **279**, 724–736.
- Thomas, S. *et al.* (2010) CAMP: a useful resource for research on antimicrobial peptides. *Nucleic Acids Res.*, **38**, 774–780.
- Tossi, A. and Sandri, L. (2002) Molecular diversity in gene-encoded, cationic antimicrobial polypeptides. *Curr. Pharm. Des.*, **8**, 743–761.
- Wang, G. *et al.* (2009) APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucleic Acids Res.*, **37**, D933–D937.
- Yang, X. *et al.* (2012) Extremely abundant antimicrobial peptides existed in the skins of nine kinds of Chinese odorous frogs. *J. Proteome Res.*, **11**, 306–319.