Current trends in antimicrobial agent research: chemo- and bioinformatics approaches

Riadh Hammami and Ismail Fliss

STELA Dairy Research Center, Nutraceuticals and Functional Foods Institute, Université Laval, G1K 7P4 Québec, QC, Canada

Databases and chemo- and bioinformatics tools that contain genomic, proteomic and functional information have become indispensable for antimicrobial drug research. The combination of chemoinformatics tools, bioinformatics tools and relational databases provides means of analyzing, linking and comparing online search results. The development of computational tools feeds on a diversity of disciplines, including mathematics, statistics, computer science, information technology and molecular biology. The computational approach to antimicrobial agent discovery and design encompasses genomics, molecular simulation and dynamics, molecular docking, structural and/or functional class prediction, and quantitative structure–activity relationships. This article reviews progress in the development of computational methods, tools and databases used for organizing and extracting biological meaning from antimicrobial research.

Introduction

The increasing resistance of dangerous microbes to conventional antibiotics has created demand for new antimicrobial agents. Computational approaches such as chemo- and bioinformatics are accelerating the process of antimicrobial drug discovery and design by providing a rational basis for the selection of chemical structure. Genomics, molecular simulation and dynamics, molecular docking, structural/functional class prediction, and quantitative structure–activity relationships (QSARs) have all benefited from the proliferation of genomic and proteomic databases and have thus become standard tools in the quest to develop novel products for treating infections. This article reviews progress in the development of computational tools and databases used for organizing and extracting biological meaning from the comparison of large collections of genomes.

Bioinformatics databases and resources for antimicrobials research

Databases and bioinformatics tools containing genomic, proteomic and functional information have become indispensable in antimicrobial drug research. Many general databases are used in this field, such as UniProt [1] and PDB [2]. Over the past decade, databases devoted to antimicrobial peptides (AMPs) have been developed, as summarized in Table 1. A total of seven antimicrobial databases are currently described (September 2009) in the Nucleic Acids Research Molecular Biology Database Collection (http://www3.oup.co.uk/nar/database/a/) [3]. Researchers are particularly interested in the special category of agents known as AMPs. These compounds are widely distributed effectors of innate host defense in plants and animals. AMPs seem to be promising candidates for supplementing and possibly replacing older antimicrobial agents that are losing their effectiveness. One of the driving forces behind the research into AMPs has been the observation that these peptides exert a broad-spectrum antimicrobial activity and that bacteria do not seem to develop resistance as easily as with conventional antibiotics [4]. Despite their diverse origins, the majority of them have common biophysical parameters that are probably essential for activity, including small molecular size, cationicity and amphipathicity. Databases AMSDb (http://www.bbcm.univ.trieste.it/~tossi/pag1.htm), APD [5], ANTIMIC [6] and CAMP [7] all cover AMP sequences from diverse origins. Some databases focus on AMPs produced by bacteria (BACTIBASE [8]), plants (PhytAMP [9]) and shrimp (PenBase [10]), and others target specific families of AMP, including defensins (Defensins [11]), cyclotides (CyBase [12]) and peptaibols (Peptaibol [13]). AMPer [14], based on hidden Markov models

Corresponding author:. Hammami, R. (riadh.hammami@fsaa.ulaval.ca)

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2010.05.002

TABLE 1

Databases and bioinformatics resources for antimicrobials research.

Database	Creation/update	Summary	URL	Refs
UniProt	–	Resource for protein sequence and annotation data, collaboration between Switzerland, UK and the USA	http://www.uniprot.org/	[1]
The Protein Data Bank	–	A worldwide repository for the processing and distribution of 3D biological macromolecular structure data, USA	http://www.rcsb.org/pdb/	[2]
AMSDb	2002, 2004	Antimicrobial Sequences Database, Italy	http://www.bbcm.univ.trieste.it/~tossi/amsdb.html	–
ANTIMIC	2004 (presently inactive)	Database of Antimicrobial Peptides, Singapore	http://research.i2r.a-star.edu.sg/Templar/DB/ANTIMIC/	[6]
APD2	2004, 2009	The Antimicrobial Peptide Database, USA	http://aps.unmc.edu/AP/main.html	[5]
CAMP	2009	Collection of Antimicrobial Peptides, India	http://www.bicnirrh.res.in/antimicrobial	[7]
APPDb	2002 (presently inactive)	Antimicrobial Peptide and Protein Database, Ireland	http://ercbinfo1.ucd.ie/APPDb/	–
AMPer	2007	A database and discovery tool for antimicrobial peptides, based on hidden Markov models and the SwissProt databank, Canada	http://marray.cmdr.ubc.ca/cgi-bin/amp.pl	[14]
Peptaibol	2003, 2004	Peptaibol Database, UK	http://peptaibol.cryst.bbk.ac.uk/home.shtml	[13]
SAPD	2002	Synthetic Antibiotic Peptides Database, Finland	http://oma.terkko.helsinki.fi:8080/~SAPD	[15]
Defensins	2007	Defensins Knowledgebase, Singapore	http://defensins.bii.a-star.edu.sg/	[11]
CyBase	2006, 2008	A database of cyclic protein sequence and structure, Australia	http://research1t.imb.uq.edu.au/cybase/	[12]
PenBase	2005	The Shrimp Penaeidin Database, France	http://penbase.immunaqua.com/	[10]
BACTIBASE	2007, 2009	A data repository of bacteriocin natural antimicrobial peptides, Tunisia	http://bactibase.pfba-lab-tun.org/	[8]
PhytAMP	2008	A database dedicated to plant antimicrobial peptides, Tunisia	http://phytamp.pfba-lab-tun.org/	[9]
RAPD	2008, 2009	A database of recombinantly-produced antimicrobial peptides, USA	http://faculty.ist.unomaha.edu/chen/rapd/	[17]
BAGEL	2006	A genome mining tool for putative bacteriocin gene clusters detection, The Netherlands	http://bioinformatics.biol.rug.nl/websoftware/bagel/	[22]
TB Database	2009	Database for tuberculosis research, USA	http://www.tbdb.org/	[21]
AMICBASE	2005, 2008	Contains information on the antimicrobial and toxicological properties of natural compounds produced by microorganisms and higher plants, John Wiley & Sons, New York, USA	http://www.wiley.com/WileyCDA/Section/id-301570.html	[16]
Novel Antibiotics DataBase	2003	Contains substances reported first in the *Journal of Antibiotics*, Japan	http://www.nih.go.jp/~jun/NADB/search.html	–
A/OL	2002, 2004	Antimicrobial compounds, applications in the food field, The Netherlands	http://www.atoapps.nl/AOLKnowledge/	–

and the SwissProt database, is a useful discovery tool for AMPs. Designed antimicrobial peptides, or DAPs (laboratory-synthesized peptide antibiotics that display broad-spectrum antibacterial activity), are catalogued in the Synthetic Antibiotic Peptides Database (SAPD) [15]. Other recent databases focus on pharmaceutical drugs and natural compounds produced by microorganisms and higher plants (e.g. The Novel Antibiotics Database, AMICBASE [16]) and on methods rather than sequence information (RAPD [17]). Figure 1 summarizes databases dealing with AMPs. Various online data sources from surveillance studies on antimicrobial resistance are now available to practitioners such as infectious disease specialists, as well as to scientists with a research interest in the field of antimicrobial resistance. Surveys of these Internet resources have been published in recent review articles (see Refs. [18,19]).

The organization of data into a unified resource ensures data quality and integrity and brings considerable benefits in terms of synergy and efficiency; however, there are several antimicrobial databases of very similar or identical scope that act in competition rather than in a network. Cross-linking between these would make their use more efficient, and researchers would benefit from such synergy. For example, cyclotides are a group of circular plant peptides that are covered at least by three specialized databases – CyBase (299) [12], KNOTTIN (78) [20] and PhytAMP (76) [9] – and other generalized databases, such as APD (126) [5], AMSDb (4) and CAMP (24) [7]. Differences in entries are numerous and caused by a

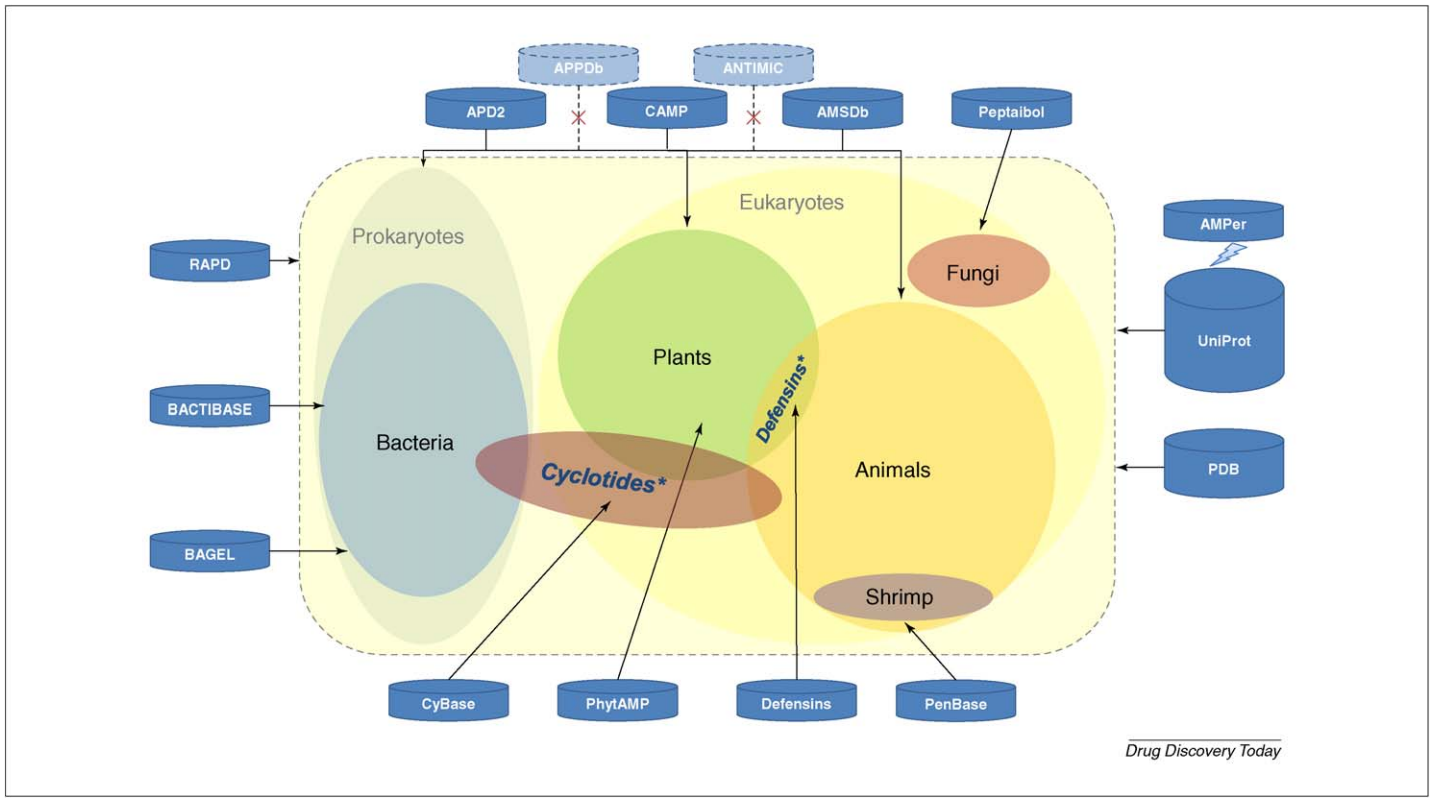

FIGURE 1

Summary of databases dealing with antimicrobial peptides (AMPs). Databases dealing with AMPs are shaped as blue cylinders. Inactive databases are grayed-out (ANTIMIC and APPDb). AMP families (cyclotides and defensins) are marked with an asterisk.

lack of updates and/or redundancies, such as sequences being present in precursor and mature forms. Sustained effort is necessary to keep a database viable and up-to-date and thereby avoid its loss, as has occurred with ANTIMIC [6] and APPDb (Table 1). During 2009, we noticed the creation of two new databases (CAMP [7] and TB database [21]) and the update of the APD [5], BACTIBASE [8] and RAPD [17] databases. Otherwise, none of the remaining databases were subject to a major update, including AMSDb, which has not been updated since 2004. In the past decade, several synthetic antibiotic peptides have been generated, and the SAPD [15] database should be updated accordingly. Table 1 provides information about the creation date and updates of these databases. Of these databases, only BAGEL [22] focuses on genomic data of bacteriocins. Conversely, all databases cover proteomic data of AMPs with different levels of curation. The Peptaibol [13] differs from the other databases with its main focus being non-ribosomally synthesized peptides isolated from fungi. Whereas APD [5] collects only 'mature and active' peptides, CAMP [7] contains a mixture of mature and precursor proteins. Compared with other databases, only CyBase [12] and PenBase [10] contain both genomic and proteomic data. AMPer [14] database could be considered as a tool for the identification of new AMPs by scanning of the SwissProt database. Some of the described databases act as simple data repositories and contain basic data about AMPs. Conversely, few databases contain various tools for sequence analysis and drug discovery. Certainly, adding analysis tools to databases enhances the way we look for information and will ultimately enable better understanding of antimicrobial agents.

Bioinformatics analysis tools and methods for antimicrobial research

The merging of mathematics, statistics, computer science, information technology and molecular biology has created an information-oriented field in biology, now known as bioinformatics. A wide variety of bioinformatics methods have been reported in the literature, including text mining, information management, sequence analysis, molecular interactions and advanced systems simulation. The development of effective computational tools depends on knowledge generated from these diverse disciplines.

Programs

Loading all of the sequence and other associated information into a relational database enables the integration of various tools encompassing homology search, multiple sequence alignment, phylogenies, physicochemical profiles, and other calculations and predictions. First, homology search engines such as BLAST [23] and FASTA [24] require the building of a formatted database. These databases can be generated for individual species genomes and proteomes, as well as the entire combined collections of DNA sequences and proteins. The homology hit is based on statistical significance (*E*-value). Once selected, sequences can be exported into programs for multiple sequence alignment and phylogenetic analyses such as CLUSTALW [25] and PHYLIP [26], respectively. The primary sequence of peptides can be analyzed using tools such as ProtParam [27] (http://www.expasy.org/tools/protparam.html) and pepstats [28] (EMBOSS package: http://www.ebi.ac.uk/emboss/). Secondary structure can be predicted using programs

such as NNPREDICT [29] (http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html) and SCRATCH servers [30] (http://www.igb.uci.edu/tools/scratch/). CyBase database provides the Diversity Wheel tool for the representation of sequence diversity from a multiple sequence alignment, positioning the consensus sequence in the inner circle with a spike protruding from each position to represent the amino acid variation observed at that position [12]. The mean hydrophobicity and relative hydrophobic moment can be calculated using the pepwheel [28] (EMBOSS package) or the CCS scale [31] (http://www.bbcm.univ.trieste.it/~tossi/HydroCalc/HydroMCalc.html). The helical wheel projection and physicochemical profiles can be calculated by various programs such as ANTHEPROT [32] and SciDBMaker [33]. Alternatively, databases APD [5], CAMP [7], PhytAMP [9] and BACTIBASE [8] offer tools for calculating physicochemical profiles of AMPs, plus predictive tools based on peptide similarity, random forests, support vector machine, discriminant analysis and hidden Markov models. Alternatively, the AntiBP server can be used to predict the antibacterial peptides in a protein sequence [34]. Prediction can be done using quantitative matrix, artificial neural network and support vector machine methods with binary patterns of peptide sequences. The predictive models used in AntiBP were developed using N- and C-terminal residues. Another predictive model, antimicrobial protein annotation, was developed by the IBM Bioinformatics group (http://cbcsrv.watson.ibm.com/Tamp.html) [35]. This tool can determine whether any parts of a peptide have the potential to exhibit antimicrobial activity. Moreover, the method can also determine which of the known AMP families is likely to be the closest to the query sequence in terms of antimicrobial behavior [35].

When the structures of experimental peptides have not been elucidated, computational methods are widely used to predict three-dimensional configuration and provide insight into their structure and function. Homology and comparative modeling are commonly used in structure-based drug discovery (for a review of this subject, see Ref. [36]). Homology and comparative modeling are used to predict the three-dimensional structure of a 'target' protein based on its sequence alignment with a related homologous protein ('template') of known structure. The accuracy of the predictions depends on the quality of the sequence alignment and template structure. High-quality structural models are produced when the target and template are closely related. The use of such methods has already proven rewarding and is increasingly widespread in innovative drug research and development. Various methods for the prediction of structure exist, including satisfaction of spatial restraints, *ab initio* fragment assembly and local similarity and/or fragment assembly. A wide range of three-dimensional structure prediction programs and web servers have been developed, including SWISS-MODEL [37], Geno3D [38] and ROBETTA [39]. MODELLER is a popular computer program for comparative protein structure modeling by satisfaction of spatial restraints [40]. In general, *ab initio* programs are used only for organic compounds and very short peptides. This method is used to predict the structures of peptides with yet unknown homologs. Programs for molecular modeling using homology are more suitable for peptides and proteins when homologs are identified. In its second release, BACTIBASE [8] database has incorporated MODELLER as tool for predicting the 3D structure of a user peptide by homology to known bacteriocins. Such a combination of relational database with molecular modeling tools should be very useful for the *in silico* design of novel AMPs.

Genomics and target discovery

Although the sequencing of bacterial genomes is becoming increasingly commonplace, identifying good antimicrobial drug targets on the basis of a genomic sequence remains a challenge. To date, the discovery of new antibiotics is still technically challenging, costly and time-consuming [41]. Despite the ever-increasing number of microbial sequencing projects, the large amount of genes with unknown biological function is still a major stumbling block in the genomics revolution. Bioinformatics methods can be used to identify homologs of known genes by comparing new sequences with sequences of biochemically characterized enzymes and proteins. Recently, Field *et al.* [42] reviewed progress in the development of computational tools and databases for the comparison of prokaryotic genomes. One aim of comparative genomics is to identify genes that encode bacterial pathogenicity factors or are essential for bacterial survival (for a review of this subject, see Refs. [43–45]). Comparisons are made between pathogenic and non-pathogenic strains, as well as pathogens versus their human and animal hosts. The latter is essential because a good antimicrobial drug target should have no homologs in mammalian cells, thus reducing the chances of undesirable side-effects. Actually, there is a growing knowledge base in host–pathogen interactions that might provide new strategies in antibiotic therapeutics. Genomics and its associated technologies are not limited to target-screening approaches and provide new insight into antimicrobial mechanisms of action (MOAs) [46]. Whole-genome expression profiling technologies are applied to the elucidation of cellular responses to treatments with antimicrobial agents [47]. Prediction of antibacterial compound MOAs by gene expression profiling is a microarray-based approach. This enables not only the screening of large numbers of natural compounds for molecules that exhibit antibacterial potential but also the classification of the MOAs of novel structural classes of antibiotics. Expression profiling of bacterial responses to antimicrobials is changing our concepts of antimicrobial–target interaction and is expected to reveal new areas for antimicrobial drug discovery [46]. Actually, various tools that handle gene profiling data are available and are expected to increase rapidly. Mandal *et al.* [48] have reviewed bioinformatics tools for gene expression analysis.

Structure-based drug design methods

QSAR methods

One of the most broadly used chemoinformatics approaches is QSAR modeling, which is widely used as a tool for antimicrobial drug discovery. QSARs can be defined as quantitative models that correlate the variation in measured biological activity with the variation in molecular structure among a series of chemical compounds. QSAR computational analyses have been applied successfully to the analysis of AMP data, resulting in a model that quantifies linear sequence patterns [49], contact energy between neighboring amino acids [50] and amphipathicity [51]. Initially, QSAR methods have been based on closely related sequence variants and typically have utilized residue-based descriptors providing only moderately useful insights. Recently, Cherkasov and

colleagues have successfully applied atomic-resolution QSARs combined with machine learning techniques, including artificial neural networks (ANNs), to a large data set of peptides containing high sequence diversity [52,53]. These works demonstrated that atomic-resolution characteristics are much more accurate for modeling of antimicrobial activity of cationic than simple considerations of polar and hydrophobic characteristics of constituent amino acids [52]. These statistical methods thus provide valuable insight into the relationships between the amino acid sequences of AMPs and their potency (for a review of the subject, see Ref. [54]). Furthermore, use of QSAR methodology combined with artificial intelligence/neural network approaches can accelerate the discovery and design of AMPs considerably. An example of a QSAR method that has been found suitable for the classification of inhibitory peptides (as antiviral/anti-HIV, anticancer/anti-tumor, antibacterial and antifungal) is increment diversity with quadratic discriminant analysis [55]. Actually, various commercial programs for QSAR computational analysis are available, including MOE (http://www.chemcomp.com/), ACD labs (http://www.acdlabs.com/) and Tsar (http://accelrys.com/products/accord/desktop/tsar.html). SYBYL-X (http://tripos.com/), another commercial software, provides a comprehensive package for structure-based design.

Artificial neural networks

Recently, ANNs have been the center of attention in the field of pattern recognition. ANNs are mathematical modeling algorithms that can offer a reliable and efficient way for *in silico* identification of novel AMPs [53]. Bucinski *et al.* [56] have demonstrated that ANN analysis could be applied successfully to the prediction of the antifungal activities of pyridine derivatives of a defined structure. ANN analysis can be performed with Tsar or Neural Network Toolbox (a MATLAB software package; http://www.mathworks.com/products/neuralnet/). Alternatively, ANN training can be conducted freely by using the neural net software package Stuttgart Neural Network Simulator (http://www.ra.cs.uni-tuebingen.de/SNNS/).

Fuzzy logic modeling

The principal objective of fuzzy (non-binary) logic modeling is the definition of a rule base that is able to predict dependent variables as a function of independent variables for some useful purpose. Mikut and Hilpert [57] recently introduced fuzzy logic with molecular descriptors for the analysis of AMPs and found it able to distinguish active and inactive peptides with impressive accuracy. The application of fuzzy methods to QSAR is expected to provide useful insight for antimicrobial drug research and development. Fuzzy calculation can be performed using Gait-CAD (http://sourceforge.net/projects/gait-cad/), an open-source MATLAB toolbox.

Molecular simulations and dynamics

AMPs interact specifically with the bacterial membrane and kill the cell by causing leakage of its contents. Although membrane interaction might be essential, alternative mechanisms have been increasingly considered as part of AMP action against microbes [58]. Different mechanisms are proposed to explain AMPs' mode of action on bacterial surfaces, including the formation of transmembrane (TM) ion-permeable pores (barrel-stave and toroidal pore models) and detergent-like binding to the membrane surface (carpet model); for a review of the subject, see Ref. [59]. One current area of research involves simulating bacterial membrane and AMP interactions with phospholipid bilayers (e.g. phosphatidylethanolamines or phosphatidylglycerol) or detergent micelles (sodium dodecylsulfate and dodecylphosphocholine). Various lipid bilayer models mimicking the Gram-positive bacterial cell membrane have been developed and are being used increasingly in antimicrobial drug research [60–62]. Matyus *et al.* [63] have provided a detailed review of computer simulation of AMPs and their interaction with lipid bilayers. There is now an abundance of freeware and commercial software for molecular dynamics (MD) simulations. CHARMM [64] (http://www.charmm.org/) is one such freeware program widely used to simulate AMP interactions with lipid bilayers or detergent micelles. Another free package, GROMACS [65] (http://www.gromacs.org/), is used in MD studies to generate 'trajectories'. Discovery Studio®, a commercial collection of software applications, provides a sophisticated solution for sequence analysis, QSAR, molecular modeling and computational simulation (http://accelrys.com/products/discovery-studio/).

MD simulations have proven particularly useful for describing the bacterial killing phenomenon and for studying AMP binding to bilayer surfaces, peptide insertion, TM domain structure and dynamics, and the properties of model pores assembled from TM aggregates. MD simulations should contribute to understanding of the structural, dynamic and functional properties of AMPs. Experimental validation using simpler models is a prerequisite for MD simulation studies of multicomponent bilayer systems. If this condition is met, the results of MD simulations might provide rational approximations applicable to real systems.

Concluding remarks

The incidence of bacterial infections in humans is becoming a major concern in both the food and the medical sectors worldwide and has created a need for novel therapeutic agents. Genome sequencing was particularly useful to understand certain mechanisms of bacterial virulence or resistance against antibacterial agents. The integration of existing structural and functional data into the study of AMPs offers an opportunity to develop models for the quantitative prediction of the effectiveness of novel peptide candidates as antimicrobial drugs. Computer-aided drug design enables the efficient analysis of manifold possibilities in a cost-effective manner by providing a reduced set of candidates to be analyzed using more capital-intensive experimental methods. Combining experimental data with computational biology will ultimately enable better understanding of antimicrobial agent–target interaction and the ability to manipulate biological systems more efficiently. To achieve this, collaboration between experimental microbiologists and bioinformaticians is necessary, and the development of databases and data-mining methods constitutes the first step of this process. The development of databases such as APD, CyBase or BACTIBASE underlies efforts to generate statistically sound QSAR models for the prediction of antimicrobial activity – models that are vital to the development of computational microbiology. The combination of bioinformatics and relational databases provides the antimicrobial researcher with better tools for analyzing, linking and comparing online search results. The development of computational tools depends on knowledge generated from diverse disciplines including mathematics, statistics, computer science, information technology

and molecular biology. Bioinformatics and wet-lab biology are interdependent and complement each other for the purposes of their own progress and for progress in antimicrobial drug discovery in the future.

Acknowledgements

The authors thank Stephen Davids for crucial reading of the manuscript. Thanks are also due to Caillard Romain for his helpful discussions.

References

1 Wu, C.H. *et al.* (2006) The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res.* 34 (Database issue), D187–D191
2 Dutta, S. *et al.* (2007) Using the tools and resources of the RCSB protein data bank. *Curr. Protoc. Bioinform.* (Chapter 1, Unit 1.9)
3 Galperin, M.Y. and Cochrane, G.R. (2009) Nucleic acids research annual database issue and the NAR online molecular biology database collection in 2009. *Nucleic Acids Res.* 37 (Database issue), D1–D4
4 Zasloff, M. (2002) Antimicrobial peptides of multicellular organisms. *Nature* 415, 389–395
5 Wang, G. *et al.* (2009) APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucleic Acids Res.* 37 (Database issue), D933–D937
6 Brahmachary, M. *et al.* (2004) ANTIMIC: a database of antimicrobial sequences. *Nucleic Acids Res.* 32 (Database issue), D586–D589
7 Thomas, S. *et al.* (2010) CAMP: a useful resource for research on antimicrobial peptides. *Nucleic Acids Res.* 38 (Database issue), D774–D780
8 Hammami, R. *et al.* (2010) BACTIBASE second release: a database and tool platform for bacteriocin characterization. *BMC Microbiol.* 10, 22
9 Hammami, R. *et al.* (2009) PhytAMP: a database dedicated to antimicrobial plant peptides. *Nucleic Acids Res.* 37 (Database issue), D963–D968
10 Gueguen, Y. *et al.* (2006) PenBase, the shrimp antimicrobial peptide penaeidin database: sequence-based classification and recommended nomenclature. *Dev. Comp. Immunol.* 30, 283–288
11 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 (Database issue), D265–D268
12 Wang, C.K. *et al.* (2008) CyBase: a database of cyclic protein sequences and structures, with applications in protein discovery and engineering. *Nucleic Acids Res.* 36 (Database issue), D206–D210
13 Whitmore, L. and Wallace, B.A. (2004) The Peptaibol Database: a database for sequences and structures of naturally occurring peptaibols. *Nucleic Acids Res.* 32 (Database issue), D593–D594
14 Fjell, C.D. *et al.* (2007) AMPer: a database and an automated discovery tool for antimicrobial peptides. *Bioinformatics* 23, 1148–1155
15 Wade, D. and Englund, J. (2002) Synthetic antibiotic peptides database. *Protein Pept. Lett.* 9, 53–57
16 Pauli, A. (2006) Anticandidal low molecular compounds from higher plants with special reference to compounds from essential oils. *Med. Res. Rev.* 26, 223–268
17 Li, Y. and Chen, Z. (2008) RAPD: a database of recombinantly-produced antimicrobial peptides. *FEMS Microbiol. Lett.* 289, 126–129
18 Falagas, M.E. and Karveli, E.A. (2006) World Wide Web resources on antimicrobial resistance. *Clin. Infect. Dis.* 43, 630–633
19 Harbarth, S. and Emonet, S. (2006) Navigating the World Wide Web in search of resources on antimicrobial resistance. *Clin. Infect. Dis.* 43, 72–78
20 Gelly, J.C. *et al.* (2004) The KNOTTIN website and database: a new information system dedicated to the knottin scaffold. *Nucleic Acids Res.* 32 (Database issue), D156–D159
21 Reddy, T.B.K. *et al.* (2009) TB database: an integrated platform for tuberculosis research. *Nucleic Acids Res.* 37 (Database issue), D499–D508
22 de Jong, A. *et al.* (2006) BAGEL: a web-based bacteriocin genome mining tool. *Nucleic Acids Res.* 34 (Web Server issue), W273–W279
23 Altschul, S.F. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402
24 Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.* 85, 2444–2448
25 Larkin, M.A. *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948
26 Felsenstein, J. (1989) PHYLIP – Phylogeny Inference Package (Version 3.2). *Cladistics* 5, 164–166
27 Gasteiger, E. *et al.* (2005) Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook* (Walker, J.M., ed.), pp. 571–607, Humana Press
28 Rice, P. *et al.* (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* 16, 276–277
29 Kneller, D.G. *et al.* (1990) Improvements in protein secondary structure prediction by an enhanced neural network. *J. Mol. Biol.* 214, 171–182
30 Cheng, J. *et al.* (2005) SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res.* 33 (Web Server issue), W72–W76
31 Tossi, A. *et al.* (2002) New consensus hydrophobicity scale extended to non-proteinogenic amino acids. *Peptides 2002: Proceedings of the Twenty-seventh European Peptide Symposium (Ziino, ed.)* pp. 416–417
32 Deleage, G. *et al.* (1988) ANTHEPROT: a package for protein sequence analysis using a microcomputer. *Comput. Appl. Biosci.* 4, 351–356
33 Hammami, R. *et al.* (2008) SciDBMaker: new software for computer-aided design of specialized biological databases. *BMC Bioinform.* 9, 121
34 Lata, S. *et al.* (2010) AntiBP2: improved version of antibacterial peptide prediction. *BMC Bioinform.* 11 (Suppl. 1), S19
35 Rigoutsos, I. and Floratos, A. (1998) Combinatorial pattern discovery in biological sequences: the TEIRESIAS algorithm. *Bioinformatics* 14, 55–67
36 Cavasotto, C.N. and Phatak, S.S. (2009) Homology modeling in drug discovery: current trends and applications. *Drug Discov. Today* 14, 676–683
37 Arnold, K. *et al.* (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22, 195–201
38 Combet, C. *et al.* (2002) Geno3D: automatic comparative molecular modelling of protein. *Bioinformatics* 18, 213–214
39 Kim, D.E. *et al.* (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* 32 (Web Server issue), W526–W531
40 Sali, A. and Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815
41 Payne, D.J. *et al.* (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* 6, 29–40
42 Field, D. *et al.* (2005) Databases and software for the comparison of prokaryotic genomes. *Microbiology* 151, 2125–2132
43 Fritz, B. and Raczniak, G.A. (2002) Bacterial genomics: potential for antimicrobial drug discovery. *BioDrugs* 16, 331–337
44 Tang, C.M. and Moxon, E.R. (2001) The impact of microbial genomics on antimicrobial drug development. *Annu. Rev. Genomics Hum. Genet.* 2, 259–269
45 Chan, P.F. *et al.* (2002) Novel antibacterials: a genomics approach to drug discovery. *Curr. Drug Targets Infect. Disord.* 2, 291–308
46 Brazas, M.D. and Hancock, R.E.W. (2005) Using microarray gene signatures to elucidate mechanisms of antibiotic action and resistance. *Drug Discov. Today* 10, 1245–1252
47 Fischer, H.-P. (2001) The impact of expression profiling technologies on antimicrobial target identification and validation. *Drug Discov. Today* 6, 1149–1150
48 Mandal, S. *et al.* (2009) Rational drug design. *Eur. J. Pharmacol.* 625, 90–100
49 Hilpert, K. *et al.* (2006) Sequence requirements and an optimization strategy for short antimicrobial peptides. *Chem. Biol.* 13, 1101–1107
50 Jenssen, H. *et al.* (2008) QSAR modeling and computer-aided design of antimicrobial peptides. *J. Pept. Sci.* 14, 110–114
51 Frecer, V. *et al.* (2004) *De novo* design of potent antimicrobial peptides. *Antimicrob. Agents Chemother.* 48, 3349–3357
52 Cherkasov, A. *et al.* (2009) Use of artificial intelligence in the design of small peptide antibiotics effective against a broad spectrum of highly antibiotic-resistant superbugs. *ACS Chem. Biol.* 4, 65–74
53 Fjell, C.D. *et al.* (2009) Identification of novel antibacterial peptides by chemoinformatics and machine learning. *J. Med. Chem.* 52, 2006–2015
54 Hilpert, K. *et al.* (2008) Short linear cationic antimicrobial peptides: screening, optimizing, and prediction. In *Peptide-Based Drug Design (Vol. 494)* (Otvos, L., ed.), pp. 127–159, Humana Press
55 Chen, W. and Luo, L. (2009) Classification of antimicrobial peptide using diversity measure with quadratic discriminant analysis. *J. Microbiol. Methods* 78, 94–96
56 Bucinski, A. *et al.* (2009) Artificial neural networks in prediction of antifungal activity of a series of pyridine derivatives against *Candida albicans*. *J. Microbiol. Methods* 76, 25–29
57 Mikut, R. and Hilpert, K. (2009) Interpretable features for the activity prediction of short antimicrobial peptides using fuzzy logic. *Int. J. Pept. Res. Ther.* 15, 129–137
58 Jenssen, H. *et al.* (2006) Peptide antimicrobial agents. *Clin. Microbiol. Rev.* 19, 491–511

59 Brogden, K.A. (2005) Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat. Rev. Microbiol.* 3, 238–250
60 Tieleman, D.P. and Sansom, M.S.P. (2001) Molecular dynamics simulations of antimicrobial peptides: from membrane binding to trans-membrane channels. *Int. J. Quantum Chem.* 83, 166–179
61 Khandelia, H. and Kaznessis, Y.N. (2007) Structure of the antimicrobial beta-hairpin peptide protegrin-1 in a DLPC lipid bilayer investigated by molecular dynamics simulation. *Biochim. Biophys. Acta* 1768, 509–520
62 Manna, M. and Mukhopadhyay, C. (2009) Cause and effect of melittin-induced pore formation: a computational approach. *Langmuir* 25, 12235–12242
63 Matyus, E. *et al.* (2007) Computer simulation of antimicrobial peptides. *Curr. Med. Chem.* 14, 2789–2798
64 Brooks, B.R. *et al.* (2009) CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30, 1545–1614
65 Hess, B. *et al.* (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4, 435–447