Bioinformatics, 31(14), 2015, 2252–2256
doi: 10.1093/bioinformatics/btv131
Advance Access Publication Date: 28 February 2015
Discovery Note

OXFORD

Sequence analysis

Mining viral proteins for antimicrobial and cell-penetrating drug delivery peptides

João Miguel Freire, Susana Almeida Dias, Luís Flores, Ana Salomé Veiga* and Miguel A.R.B. Castanho*

Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, 1649-028 Lisbon, Portugal

*To whom correspondence should be addressed.

Associate Editor: Burkhard Rost

Received on September 30, 2014; revised on February 17, 2015; accepted on February 25, 2015

Abstract

Motivation: The need for more effective and safer pharmaceuticals is a persistent quest. Microbial adaptations create the need to permanently develop new antimicrobials (AMPs), for instance. Similarly, intracellular delivery of drugs is still a challenge and translocation of membranes for drug delivery is an area of intense research. Peptides can be used both as AMP drug leads and drug carrier systems for intracellular delivery. Multifunctional proteins are abundant in viruses but, surprisingly, have never been thoroughly screened for bioactive peptide sequences.

Results: Using the AMPA and CellPPD online tools, we have evaluated the propensity of viral proteins to comprise AMP or cell-penetrating peptides (CPPs). Capsid proteins from both enveloped and non-enveloped viruses, and membrane and envelope proteins from enveloped viruses, in a total of 272 proteins from 133 viruses, were screened to detect the presence of potential AMP and CPP sequences. A pool of 2444 and 426 CPP and AMP sequences, respectively, were discovered. The capsids of flaviviruses are the best sources of these peptides reaching more than 80% of CPP sequence coverage per protein. Selected sequences were tested experimentally and validated the results. Overall, this study reveals that viruses form a natural multivalent biotechnological platform still underexplored in drug discovery and the heterogeneous abundance of CPP/AMP sequences among viral families opens new avenues in viral biology research.

Contacts: aveiga@medicina.ulisboa.pt or macastanho@medicina.ulisboa.pt

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

1 Introduction

Viruses are molecular machines that are present in all sorts of environments and conditions, being the most abundant entity in the biological world (Arias, 2013). Traditionally considered as a threat to most living systems, viruses are now seen also as an asset in biotechnological and biomedical research. Taking advantage of their natural properties, such as encapsulation of materials, self-packing, cell internalization by diverse routes and multivalent post-translational and chemical modifications to display surface ligands, viruses are now important tools in the fields of nanoengineering and nanomedicine. For instance, viral-like particles are used for gene delivery and viral capsid nanocapsules to incorporate bioactive compounds (Douglas and Young, 2006; Usme-Ciro *et al.*, 2013).

There is a constant endless quest from the scientific and pharmaceutical communities to find innovative approaches to overcome issues on safety or efficacy of current drugs, such as the increased microbial resistance to conventional antibiotics (Fox, 2013) or the inability of most therapeutic molecules to reach their cellular targets (Copolovici *et al.*, 2014; de Figueiredo *et al.*, 2014). Peptide and protein-based drugs are good drug leads because they can find use as

antimicrobial antimicrobial agents and/or drug delivery systems (Copolovici *et al.*, 2014; Fox, 2013; Koren and Torchilin, 2012). Antimicrobial peptides (AMP, Melo *et al.*, 2009) and cell-penetrating peptides (CPPs, de Figueiredo *et al.*, 2014) are two classes of peptides of which biomedical applications are expanding (Copolovici *et al.*, 2014; Fox, 2013). AMP (Fox, 2013; Zasloff, 2002) and CPP (de Figueiredo *et al.*, 2014) are typically short cationic peptides, usually 5–30 amino acid residues long, that are found in a wide diversity of natural sources from microbes to plants and animals (Milletti, 2012). AMP and CPP sequences may result from the identification of bioactive compounds on natural extracts, *in silico* analysis of natural proteins, de novo or structure-based design or chimeras of peptide fragments (Fig. 1 and Supplementary Table S1). Concerning AMP and CPP from natural origin, Figure 1 shows that only a few of them, identified up to now, rely on viral protein sources. Albeit rare, viral-based CPP are more frequent than AMP, which is in agreement with data in the literature (Douglas and Young, 2006; Usme-Ciro *et al.*, 2013). Interestingly, the first CPP that was discovered was from a viral source; Frankel and Pabo (1988) found that the HIV-1 transcription-transactivating protein (Tat) could enter cells and translocate into the nucleus. In addition to Tat, there were CPP found in proteins from other viruses such as herpes simplex (Falanga *et al.*, 2011), herpesvirus 8 (Gautam *et al.*, 2012), human respiratory syncytial virus (Langedijk, 2002), hepatitis-B (Montrose *et al.*, 2013), flock house virus (Nakase *et al.*, 2009) and alphaviruses (Langedijk *et al.*, 2004).

More recently, our group characterized Dengue virus capsid protein (DENV C) as a supercharged protein (SCP) (Freire *et al.*, 2013), with ability to efficiently deliver nucleic acids into a wide range of cells.

Two intrinsic DENV C domains with CPP properties, pepM and pepR, were identified (Freire *et al.*, 2014) (Supplementary Fig. S2A). Moreover, all *Flavivirus* capsid proteins revealed to belong to the SCP family suggesting that these proteins contain intrinsic cell-penetrating domains (Freire *et al.*, 2013).

Despite being more rare, AMP have also been found in viral proteins. AMP sequences were found in proteins of hepatitis B virus, Bacillus phage SPbeta, Bean golden mosaic virus, Lactococcus phage mv1, Pneumococcus phage Dp-1 and Streptococcus phage Cp-1 (Waghu *et al.*, 2014).

The rapidly expanding primary data volume on both AMP and CPP raised the need to create systematic databases with their sequences, origin, activity and mode-of-action (Supplementary Table S1 shows a summary of the available repositories on AMP and CPP). It is noticeable that the information regarding CPP is confined to CPPsite (Gautam *et al.*, 2012), whereas for AMP, several databases exist. A search in CAMP (Waghu *et al.*, 2014) and CPPsite (Gautam *et al.*, 2012) databases shows that about 46% and 63% of the known AMP and CPP, respectively, are from natural origin (Fig. 1). From those, only 0.1% of the AMP and 11% of the CPP are from viral source. In this study, we performed a thorough search for the discovery of AMP and CPP in all structural viral proteins available in the Viralzone database (Hulo *et al.*, 2011) using AMPA (Torrent *et al.*, 2012) and CellPPD (Gautam *et al.*, 2013) servers, respectively (Supplementary Fig. S1). An AMP/CPP sequence heat map was thus obtained for each protein. The corresponding AMP and CPP sequences are libraries of potential antibiotic and drug delivery systems drug leads, respectively, to be further explored. As a basic proof-of-principle validation of viral AMP and CPP discovery, we screened DENV C protein for AMP (Alves *et al.*, 2010) and CPP (Freire *et al.*, 2013, 2014) (Supplementary Fig. S2).

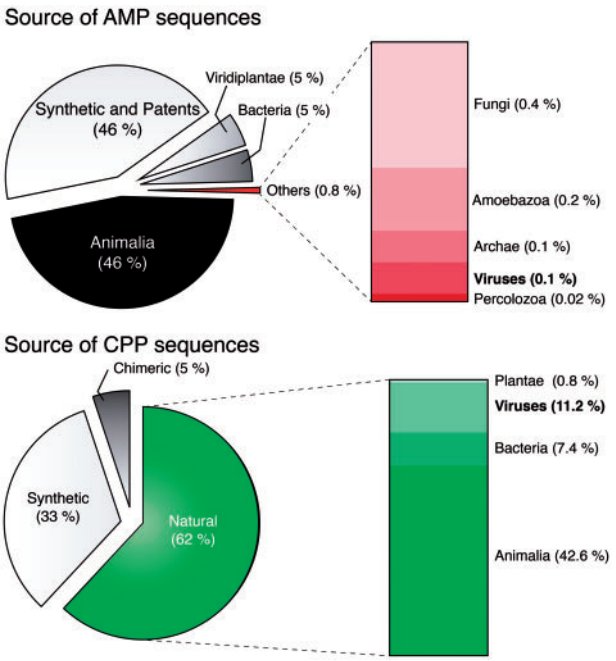

Fig. 1. Origin of AMP/CPP sequences present in databases. Top panel—Relative frequency of the origin of AMP sequences annotated at the CAMP AMP database (Waghu *et al.*, 2014). Bottom panel—Relative frequency of the origin of CPP sequences annotated at the CPPsite (Gautam *et al.*, 2012) (the databases were consulted in January 2014). (Color version of this figure is available at *Bioinformatics* online.)

2 Meta-analysis of viral protein repositories for AMP and CPP

The study of protein domains is highly relevant in biochemistry and molecular biology. Researchers aim to unravel and catalog the biological function of sequences of amino acid residues. Several repositories have been created to assign, curate and predict the structure and function of protein domains (Bateman *et al.*, 1999; Marchler-Bauer *et al.*, 2010; The UniProt Consortium, 2009; Wilkins *et al.*, 1999). These domains vary from regions prone to post-translational modifications (Bateman *et al.*, 1999; Hulo *et al.*, 2008) and catalytic sites in enzymes (Hulo *et al.*, 2008; Wilkins *et al.*, 1999) to sequence signal peptides (Hiller *et al.*, 2004). More specific domains, such as AMP/lytic domains, present in bacterial proteins bacteriocins (de Jong *et al.*, 2010; Hammami *et al.*, 2010), and cell-penetrating protein transduction domains (van den Berg and Dowdy, 2011; Zahid and Robbins, 2012) are also annotated. This article deals with the discovery of AMP and CPP sequences within viral proteins, which revealed to be a rich source for novel drugs.

The amino acid residue sequences of the structural proteins of viruses were analyzed to detect the domains with intrinsic propensity to constitute AMP or CPP. The flowchart in Supplementary Figure S1 explains the main steps for AMP/CPP screening in viral proteins. First, the 272 annotated proteins from 133 viruses in ViralZone (Hulo *et al.*, 2011) were classified according to the Baltimore system, which clusters viruses according to their type of genome. Furthermore, the proteins were divided in three different classes: Capsid (C), Envelope (E) and Membrane (M). Each protein sequence was obtained from the UniProt database (The UniProt Consortium, 2009). All the sequences were then submitted to CellPPD (Gautam *et al.*, 2013) for CPP identification and AMPA (Torrent *et al.*, 2012) to predict AMP sequences. A library of 2444 potential new CPP and a library of 426 potential new AMP were

obtained. Supplementary Section 2 and Viral_proteins.xls provide more details regarding the predicting tools parameters and analysis as well as the viral proteins used in the query.

We used DENV C as a model protein to validate the meta-analysis described above. AMPA retrieved the sequence 66-99 as AMP and CellPPD identified the entire sequence of the protein as having CPP properties (Supplementary Fig. S2A). The peptide which sequence corresponds to amino acid residues 67–100 from DENVC is known as pepR and has been characterized experimentally as an antibiotic (Alves *et al.*, 2010) (Supplementary Fig. S2B). The whole protein has cell-penetrating ability, which is characteristic of SCP. This ability is maintained by the individual domains of the protein (Freire *et al.*, 2013, 2014) (Supplementary Fig. S2). In addition, two other proteins with well-characterized AMP and CPP domains were also used to validate the methodology: HIV-Tat and the Human neutrophil Bactericidal/permeability-increasing protein (BPI) (Supplementary Fig. S3). HIV-Tat is the template for one of the 'gold standard' CPP: tat (Frankel and Pabo, 1988; Vivès *et al.*, 1997). The protein region comprised by amino acid residues 37–72 was targeted by CellPPD and includes the tat sequence (amino acid residues 37-60). AMPA revealed sequences 23-44 and 47-62 as intrinsic AMP motifs, in agreement with experimental data (Splith and Neundorf, 2011).

AMPA also retrieved the BPI AMP domains, corresponding to the sequences 84–101 and 142–162, which are included in the 21 kDa N-terminal fragment of rBPI21. This fragment is known for the strong AMP activity (Domingues *et al.*, 2009, 2014).

CellPPD is a SVM tool trained with several proteins including HIV-Tat but not BPI or DENV C. The CPP sequence found in Tat reveals the self-consistency of the method. The retrieval of AMP and CPP validated experimentally in these proteins demonstrates the robustness of the method to identify positive hits. The possibility of false positives in the libraries of new CPP and AMP, however, remained and had to be addressed experimentally. Fourteen CPP sequences and 15 AMP sequences were selected to be tested experimentally for their efficacy. A broad range of conditions (hydrophobicity, amphipathicity and hit scores) were covered (Supplementary Information, Section 4.1, Supplementary Table S2 and Fig. S8). Cell penetration in a human cell line and antibacterial activity against two bacterial strains reveal that very few false positives exist for CPP, in contrast with AMP. AMP found range from moderate to extremely potent, but the fraction of false positives is significant.

CellPPD retrieved extensive putative CPP regions in C proteins of flaviviruses. AMP profiling, despite less consistent, is still significantly shared over these proteins (Table 1 and Supplementary Fig. S5). This homogeneity is not surprising because all documented flaviviruses C proteins are SCP (Freire *et al.*, 2013). The same evaluation was carried out for flaviviruses E and M structural proteins (Table 1, Supplementary Figs. S6 and S7). These proteins share higher sequence conservation than C proteins (Ma *et al.*, 2004). The search for AMP and CPP sequences within M proteins revealed that these proteins lack AMP sequences and are poor in CPP sequences. As for E proteins, the results show that within the sequences, the region between amino acid residues 100 and 250 has high propensity to be a CPP. This result is consistent with the fact that E proteins are involved in promoting viral-host membrane hemifusion during viral infection (Pierson and Kielian, 2013).

Table 1 summarizes the data of the occurrence of intrinsic AMP and/or CPP sequences in each class of structural proteins from flaviviruses. C proteins clearly show high potential as sources for peptide sequences with cell-penetrating properties. Moreover, data in

Table 1. Fraction of AMP and cell-penetrating domains in structural viral proteins

	%CPP			%AMP		
	Capsid	Membrane	Envelope	Capsid	Membrane	Envelope
Flaviviruses	82.4	5.7	4.1	12.3	0.0	4.5
ssRNA (+)	26.9	11.6	11.8	12.0	6.7	7.7
dsDNA	12.8	13.7	12.5	7.7	4.9	7.2
ssDNA	24.3	3.1	0.0	19.1	0.0	0.0
dsRNA	6.8	1.7	8.9	2.8	1.7	4.8
ssRNA (−)	14.3	7.1	7.3	3.5	3.5	8.5
dsDNA(RT)	24.4	10.0	15.3	12.1	6.7	7.4
ssRNA(RT)	23.3	0.0	12.6	17.3	0.0	1.9

ss, single-stranded; ds, double-stranded; (+), positive sense; (−), negative sense; RT, reverse transcribing. The percentage of amino acid residues present in CPP (%CPP) or AMP (%AMP) domains was calculated as detailed in Section 2 of Supplementary Information for capsid, envelope and membrane proteins from flaviviruses and viral classes according to the Baltimore classification as cataloged in Viralzone (Hulo *et al.*, 2011). Only viruses with characterized structural protein (capsid, envelope and membrane) sequences in UniProt were selected for the analysis. Background colors highlight different intervals: [0.0, 10.0], [10.0, 20.0], [20.0, 30.0] and >30.0.

Table 1 are evidence that C and E proteins are important sources of CPP and AMP in all viruses, not only flaviviruses. It shows in detail the propensity of finding AMP and CPP sequences in C, E and M viral proteins according to Baltimore's viral families. ssRNA(+) [e.g. *Flaviridae* (Hepatitis C, DENV, West Nile), *Togaviridae* (Sindbis virus, Chinkungunya), *Picornaviridae* (Poliovirus, Foot-and-mouth disease virus)] and double-stranded DNA(RT) [e.g. *Hepadnaviridae* (Hepatitis B, Cauliflower mosaic virus)] are the families with the biggest relative abundance of CPP or AMP domains. It is important to stress that all families have proteins with a significant extent of CPP domains, ranging from nearly 9% in the envelope proteins of double-stranded RNA to more than 30% in the C proteins of ssRNA(+).

The data collected allowed identifying 426 and 2444 new viruses-derived AMP and CPP, respectively. The similarity among the new AMP is nil (Supplementary Information, Section 2.3). Although there is mild similarity among new CPP, both libraries are remarkable considering that they are the result of guided rational drug discovery. Viruses encode, package and transport viral nucleic acids. It is, therefore, not surprising that they contain proteins with domains that are membrane active (AMP or CPP). The unusually high charge density of viral C proteins, compatible with its inherent genome binding and delivery functions is the basis for the huge potential these molecules have to include AMP and CPP sequences. Nevertheless, the incidence of CPP on specific protein classes with preferred viral families shows this property fits specific biological functions.

Viruses form functional macromolecular assemblies of proteins, genomic data (RNA or DNA) and in some cases lipids, which are geometrically organized and very elegantly built (Abrescia *et al.*, 2012). These entities self-assemble forming metastable structures, interact with membranes (host and viral), while transporting nucleic acids. Multifunctional proteins confer greater evolutionary fitness to a virus with length-wise restrained genomes, a strategy also shared in some bacteria such as *E. coli* (Ouzounis, 2000). Completely adapted to surpass different chemical and cellular barriers to fulfill its ultimate goal, viruses are in general composed by very few genes (Ignacio-Espinoza *et al.*, 2013). Thus, viral proteins must be

multifunctional entities in order to be able to perform diverse tasks during viral infection.

3 Conclusions

The data presented in this study shows that structural viral proteins are rich scaffolds for new AMP and CPP. Moreover, new biological questions are raised regarding the ubiquity of AMP and cell-penetrating properties of viral protein domains, with special incidence in capsid proteins of specific viral families, such as ssDNA, ssRNA(RT) and ssRNA(+).

Identifying potential AMP sequences from structural viral proteins is a new tool available to biotech- and pharma companies in their quest for new leads of AMP agents. In addition to the amino acid sequences identified directly by the methodologies applied in our study, the possibility of constructing chimeric drug leads makes viral proteins-associated drug discovery and delivery even more appealing. Because of their remarkable properties, the number of peptide-based drugs and drug candidates have been steadily increasing in the past decade (Kaspar and Reichert, 2013), with more than 100 having reached the market and some already surpassed the global sales threshold of a thousand million US$ (Craik *et al.*, 2012). CPP and AMP are key players in the development of innovative drug delivery systems (Huang *et al.*, 2013) and reliable alternatives to conventional antibiotics (Fox, 2013).

We have turned a corner from viewing viruses only as hostile enemies to using them as an asset for the discovery of new AMP and CPP leads, a research and development field restlessly seeking for new ideas and formulations. The usage of these natural occurring and evolutionarily optimized viral-encrusted AMP and CPP sequences paves the way for a new source of biotechnological leads, enabling more rational/guided molecular designs, which may impact directly on the advance of biomedical areas such as genetic manipulation. Moreover, chemical grafting of the viral AMP and CPP to nanoparticles and nanotechnologies may be of paramount importance to modulate their properties, combining the best of evolution and truly intelligent design. Nevertheless, it should be stressed that the method by itself does not assure that the newly found CPP will perform better in biomedical applications than existing ones. Regarding AMP, it is remarkable that the method has retrieved an AMP sequence having minimum inhibitory concentration (MIC) < 1 μM. This high activity is extremely rare for membrane-targeting AMP (Melo *et al.*, 2009). Finding such an effective AMP is not only astonishing, it also proves that both external and internal membranes in bacteria can be disrupted with similar mechanisms, a controversial but evidence-based hypothesis (Freire *et al.*, 2015).

Funding

This work was supported by Fundação para a Ciência e Tecnologia—Ministério da Educação e Ciência (FCT-MEC, Portugal) [HIVERA/0002/2013]. J.M.F. and A.S.V. acknowledge FCT-MEC for PhD fellowship SFRH/BD/70423/2010 and fellowship IF/00803/2012 under the FCT Investigator Programme.

Conflict of Interest: none declared.

References

Abrescia,N.G.A. *et al.* (2012) Structure unifies the viral universe. *Annu. Rev. Biochem.*, **81**, 795–822.

Alves,C.S. *et al.* (2010) *Escherichia coli* cell surface perturbation and disruption induced by antimicrobial peptides BP100 and pepR. *J. Biol. Chem.*, **285**, 27536–27544.

Arias,C.F. (2013) Virus diversity and evolution. *Curr. Opin. Microbiol.*, **16**, 465–467.

Bateman,A. *et al.* (1999) Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins. *Nucleic Acids Res.*, **27**, 260–262.

Copolovici,D.M. *et al.* (2014) Cell-penetrating peptides: design, synthesis, and applications. *ACS Nano*, **8**, 1972–1994.

Craik,D.J. *et al.* (2012) The future of peptide-based drugs. *Chem. Biol. Drug Des.*, **81**, 136–147.

de Figueiredo,I.R. *et al.* (2014) Cell penetrating peptides: A tool for effective delivery in gene-targeted therapies, *IUBMB Life*.doi:10.1002/iub.1257.

de Jong,A. *et al.* (2010) BAGEL2: mining for bacteriocins in genomic data. *Nucleic Acids Res.*, **38**, W647–W651.

Domingues,M.M. *et al.* (2009) rBPI21 promotes lipopolysaccharide aggregation and exerts its antimicrobial effects by (hemi)fusion of PG-containing membranes. *PLoS One*, **4**, e8385.

Domingues,M.M. *et al.* (2014) Antimicrobial protein rBPI21-induced surface changes on Gram-negative and Gram-positive bacteria. *Nanomedicine*, **10**, 543–551.

Douglas,T. and Young,M. (2006) Viruses: making friends with old foes. *Science*, **312**, 873–875.

Falanga,A. *et al.* (2011) A peptide derived from herpes simplex virus type 1 glycoprotein H: membrane translocation and applications to the delivery of quantum dots. *Nanomedicine*, **7**, 925–934.

Fox,J.L. (2013) Antimicrobial peptides stage a comeback. *Nat. Biotechnol.*, **31**, 379–382.

Frankel,A.D. and Pabo,C.O. (1988) Cellular uptake of the tat protein from human immunodeficiency virus. *Cell*, **55**, 1189–1193.

Freire,J.M. *et al.* (2013) Intracellular nucleic acid delivery by the supercharged dengue virus capsid protein. *PLoS One*, **8**, e81450.

Freire,J.M. *et al.* (2014) Nucleic acid delivery by cell penetrating peptides derived from dengue virus capsid protein: design and mechanism of action. *FEBS J.*, **281**, 191–215.

Freire,J.M. *et al.* (2015) Shifting gear in antimicrobial and anticancer peptides biophysical studies: from vesicles to cells. *J. Pept. Sci.*, **21**, 178–185.

Gautam,A. *et al.* (2012) CPPsite: a curated database of cell penetrating peptides. *Database*, **2012**, bas015.

Gautam,A. *et al.* (2013) In silico approaches for designing highly effective cell penetrating peptides. *J. Transl. Med.*, **11**, 74.

Hammami,R. *et al.* (2010) BACTIBASE second release: a database and tool platform for bacteriocin characterization. *BMC Microbiol.*, **10**, 22.

Hiller,K. *et al.* (2004) PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res.*, **32**, W375–W379.

Huang,Y. *et al.* (2013) Curb challenges of the 'Trojan Horse' approach: smart strategies in achieving effective yet safe cell-penetrating peptide-based drug delivery. *Adv. Drug Deliv. Rev.*, **65**, 1299–1315.

Hulo,N. *et al.* (2008) The 20 years of PROSITE. *Nucleic Acids Res.*, **36**, D245–D249.

Hulo,C. *et al.* (2011) ViralZone: a knowledge resource to understand virus diversity. *Nucleic Acids Res.*, **39**, D576–D582.

Ignacio-Espinoza,J.C. *et al.* (2013) The global virome: not as big as we thought? *Curr. Opin. Virol.*, **3**, 566–571.

Kaspar,A.A. and Reichert,J.M. (2013) Future directions for peptide therapeutics development. *Drug Discov. Today*, **18**, 807–817.

Koren,E. and Torchilin,V.P. (2012) Cell-penetrating peptides: breaking through to the other side. *Trends Mol. Med.*, **18**, 385–393.

Langedijk,J.P.M. (2002) Translocation activity of C-terminal domain of pestivirus Erns and ribotoxin L3 loop. *J. Biol. Chem.*, **277**, 5308–5314.

Langedijk,J.P.M. *et al.* (2004) New transport peptides broaden the horizon of applications for peptidic pharmaceuticals. *Mol. Divers.*, **8**, 101–111.

Ma,L. *et al.* (2004) Solution structure of dengue virus capsid protein reveals another fold. *Proc. Natl Acad. Sci. USA*, **101**, 3414–3419.

Marchler-Bauer,A. *et al.* (2010) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.*, **39**, D225–D229.

Melo,M.N. *et al.* (2009) Antimicrobial peptides: linking partition, activity and high membrane-bound concentrations. *Nat. Rev. Microbiol.*, **7**, 245–250.

Milletti,F. (2012) Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discov. Today*, **17**, 850–860.

Montrose,K. *et al.* (2013) Xentry, a new class of cell-penetrating peptide uniquely equipped for delivery of drugs. *Sci. Rep.*, **3**, 1661.

Nakase,I. *et al.* (2009) Cell-surface accumulation of flock house virus-derived peptide leads to efficient internalization via macropinocytosis. *Mol. Ther.*, **17**, 1868–1876.

Ouzounis,C.A. (2000) Global properties of the metabolic map of *Escherichia coli*. *Genome Res.*, **10**, 568–576.

Pierson,T.C. and Kielian,M. (2013) Flaviviruses: braking the entering. *Curr. Opin. Virol.*, **3**, 3–12.

Splith,K. and Neundorf,I. (2011) Antimicrobial peptides with cell-penetrating peptide properties and vice versa. *Eur. Biophys. J.*, **40**, 387–397.

The UniProt Consortium (2009) The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res.*, **38**, D142–D148.

Torrent,M. *et al.* (2012) AMPA: an automated web server for prediction of protein antimicrobial regions. *Bioinformatics*, **28**, 130–131.

Usme-Ciro,J.A. *et al.* (2013) Cytoplasmic RNA viruses as potential vehicles for the delivery of therapeutic small RNAs. *Virol. J.*, **10**, 185.

van den Berg,A. and Dowdy,S.F. (2011) Protein transduction domain delivery of therapeutic macromolecules. *Curr. Opin. Biotechnol.*, **22**, 888–893.

Vivès,E. *et al.* (1997) A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J. Biol. Chem.*, **272**, 16010–16017.

Waghu,F.H. *et al.* (2014) CAMP: collection of sequences and structures of antimicrobial peptides. *Nucleic Acids Res.*, **42**, D1154–D1158.

Wilkins,M.R. *et al.* (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.*, **112**, 531–552.

Zahid,M. and Robbins,P.D. (2012) Protein transduction domains: applications for molecular medicine. *Curr. Gene Ther.*, **12**, 374–380.

Zasloff,M. (2002) Antimicrobial peptides of multicellular organisms. *Nature*, **415**, 389–395.