Universidade Católica Portuguesa
Faculdade de Engenharia

Membrane-active Peptides from Structural Viral Proteins:

Identifying Novel Delivery Vectors for Gene therapy

Luís Rafael Pereira do Carmo Flores

Dissertation to obtain the Master of Science Degree in Biomedical Engineering: Specialization in Biomolecular, Tissue and Organ Engineering

Jury

Professor Doctor Manuel José Martinho Barata Marques (President)
Doctor Fábio Monteiro Fernandes
Professor Doctor Cecília Ribeiro da Cruz Calado
Professor Doctor Miguel Augusto Rico Botas Castanho (Supervisor)

March 2014

Resumo

Péptidos activos em membranas são relevantes em diversos campos da biomedicina. Os péptidos translocadores de membranas (CPPs), em particular, são promissores na administração de fármacos, incluindo em terapia génica.

O presente trabalho teve por objectivo identificar novas sequências CPPs em proteínas estruturais de vírus utilizando técnicas bioinformáticas e validação experimental. 270 proteínas virais foram examinadas para reconhecimento de potenciais CPPs, tendo sido identificadas 2400 sequências putativas. 14 CPPs de vírus foram seleccionados para ensaios *in vitro* como vectores para carga génica, utilizando oligonucleótidos de ssDNA como modelo. A eficiência de entrega foi monitorizada por espectroscopia de fluorescência, citometria de fluxo e microscopia confocal. Adicionalmente, efectuaram-se ensaios biofísicos para compreender propriedades físico-químicas necessárias para entrega celular eficiente dos CPPs. Consequentemente, medidas de potencial de membrana com di-8-ANEPPS foram utilizados ao estudar afinidade de CPPs para membranas. Foi usado dicroísmo circular para inferir estruturas secundárias induzidas em CPPs por membranas lipídicas. A conjugação entre CPPs de vírus e oligonucleótidos foi também avaliada por dispersão dinâmica de luz para aferir a formação de complexos entre vectores e carga transportada.

Seis dos péptidos demonstraram eficiência na entrega de ssDNA a células. Dados biofísicos demonstraram que a eficiência da entrega de CPPs está dependente das interacções entre CPPs e lípidos, assim como da capacidade de conjugação com a carga a transportar. Dois CPPs foram particularmente eficientes e deverão continuar sob desenvolvimento e caracterização.

Proteínas estruturais de vírus são uma fonte viável de CPPs, e podem ser exploradas para outras biotecnologias de péptidos activos em membranas, nomeadamente péptidos antimicrobianos.

Palavras-chave: Proteínas estruturais de vírus, Péptidos translocadores de membranas, Administração controlada de fármacos, Terapia génica, Péptidos activos em membranas, Bioquímica física

Abstract

Membrane-active peptides provide wide therapeutic potential in several biomedical areas. Among these, cell-penetrating peptides (CPPs) are highly promising molecules in drug delivery, particularly when applied to gene therapy applications.

This work aimed to identify novel CPP sequences in structural viral proteins using bioinformatics, followed by experimental validation. 270 structural viral proteins were screened for the existence of potential CPP sequences, which resulted in the identification of 2400 putative sequences. A subset of 14 viral CPPs was selected for *in vitro* testing as gene cargo vectors using a 15-mer ssDNA oligonucleotide as a model. Delivery efficiency was monitored by fluorescence spectroscopy, flow cytometry and confocal microscopy. Furthermore, biophysical assays were conducted to understand the physical-chemical properties required for effective CPP cellular delivery. As such, membrane dipole potential sensing, using di-8-ANEPPS, was employed to study the affinity of CPPs towards lipid membranes. Circular dichroism was used to infer about lipid membrane-induced CPP secondary structure. Moreover, conjugation between each viral CPP and oligonucleotides was evaluated by dynamic light scattering to infer about the proper formation of vector:cargo complexes.

Six peptides demonstrated clear efficiency in delivering ssDNA into cells. Biophysical data showed that the molecular determinants required for an efficient CPP are dependent on CPP-lipid interactions and proper conjugation with the cargo to deliver. Thus, two CPPs were particularly efficient and should be considered for future development and characterization.

Structural viral proteins are a viable source for new CPPs, which may also be explored for other membrane-active peptide biotechnologies, namely antimicrobial peptides.

Keywords: Structural viral proteins, Cell-penetrating peptides, Drug delivery, Gene therapy, Membrane-active peptides, Physical biochemistry

Acknowledgements

I would like to start by thanking Professor Miguel Castanho for welcoming me into his laboratory. The Professor is an excellent supervisor at a professional and personal level, and I have been most fortunate to be able to conduct research under his inspirational guidance.

I must equally acknowledge Dr. Ana Salomé Veiga for the long hours she has put into helping me perfect this work. I have learned important lessons from her, for which I am very grateful.

A special acknowledgment is due to João Freire, who has introduced me into the exciting field of cell-penetrating peptides. He has taught me how to be an independent worker and his support has never faltered when needed. João is a dedicated individual, relentless in his research which he conducts with proficiency and passion. I am certain he has a very bright future in the academic world and it has been a wonderful opportunity to work alongside such a model student.

To Tiago Figueira, Sandra Pinto and Diana Gaspar, as well as to the remaining members of MCastanho lab I leave my best regards. They have been great colleagues, very welcoming, and although excellent professionals, they have made the office a relaxed environment and a great place to work in.

To the members of Faculdade de Engenharia da Universidade Católica Portuguesa, former professors and peers alike, I express my gratitude. You have been part an integral part of my growth as a student.

Finally, to my parents and family who have always supported me and who I deeply love and respect, and to Raquel Vaz, to whom I owe so much, I dedicate this dissertation to you. All the setbacks and strife I could not have surpassed if not for your loving support and wise words, for which I am truly thankful.

To all those in my life, you have never been any less than excellent, and I sincerely hope this work does justice to all I've learned with you during the past five years, and that it is but a small sample of all there is still to come.

Contents

Figure List

Table List

Acronyms and symbols

($\Delta\psi$) - Transmembrane potential

(ψ_s) - Surface potential

(ψ_d) - Dipole potential

ΔA –Difference in absorption of circularly polarized light components

θ - Ellipticity

A-488 – Alexa Fluor 488

AMP - Antimicrobial peptide

BHK 21 – Baby hamster kidney 21

C – Peptide concentration

CD – Circular dichroism

CM – Confocal laser scanning microscopy

CPP – Cell-penetrating peptide

d - Optical path of light

di-8-ANEPPS – 4-[2-[6-(dioctylamino)-2-naphthalenyl]ethenyl]-1-(3- sulfopropyl)- pyridinium

D – Diffusion constant

DDS - Drug Delivery System

DENV – Dengue virus

DLS – Dynamic light scattering

FBS – Fetal bovine serum

FC - Flow cytometry

FS – Fluorescence spectroscopy

GFP - Green fluorescent protein

HEK 293 – Human embryonic kidney 293

HEPES - N-2-hydroxyethylpiperazine-N′-2-ethanesulfonic acid

Hr – Hydrodinamic radius

LUV - Large unillamelar vesicle

MAP – Membrane-active peptide

Mr – Mean residue molecular weight

NLS – Nuclear location signal

NP – Nanoparticle

P:L – Peptide-to-lipid ratio

PBS – Phosphate buffer saline

PC - Phosphatidylcholine

PE - phosphatidyletanolamine

PMO - Phosphorodiamidate morpholino oligomer

PNA -Peptide nucleic acid

POPC - 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

POPS - 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine

PR – Microplate reader

PS - Phosphatidylserine

R - Ratio of fluorescence intensities

$\mathbf{R}_0$ – Ratio of fluorescence intensities in undisturbed di-8-ANEPPS LUV spectrum

ssDNA - Single Strand DNA

ssDNA-A488 – ssDNA oligonucleotides labeled with A-488

ssRNA - Single Strand RNA

SCP - Supercharged protein

SEM – Standard error of the mean

SUV – Small unillamelar vesicle

SVM – Support vector machine

TAT – HIV Trans-Activator of Transcription protein

Tat – CPP derived from TAT

viralCPP – Putative CPP identified in a structural viral protein

viralCPP:ssDNA-A488 - Electrostatically bound conjugates of viralCPP and ssDNA-A488

1. Introduction

1.1. Drug delivery systems

The administration of classical drugs, subscribed to Lipinski's rule of five, as well as of state-of-the art therapeutics, such as protein/peptide or gene therapies, are a major aspect of medical care. However, these treatments are often associated with undesirable outcomes. For instance, adverse off-target effects can result from an inherent lack of specificity of pharmaceutical molecules, and are associated with high dosages and toxicity (1). Additional downsides of these drugs include systemic degradation and clearance by the renal and reticulo-endothelial systems, poor solubility, stability, circulation time, and bioavailability/biodistribution *in vivo* (2-4). Furthermore, macromolecules cannot transpose the cytoplasmic membrane of cells, which is a limiting step in the development of biomolecular therapies aimed at intracellular targets (5, 6). This is particularly relevant as it has been estimated that only 10% of the known drugable genome can be targeted by conventional drugs (7). Altogether, these issues greatly diminish pharmacological potential and are major hindrances to the pharmaceutical and biotechnological industries (2).

In this context, the emergence of improved drug delivery systems (DDSs) is expected to be of seminal importance by providing tools to overcome such limitations. Drug delivery can be defined as a multidisciplinary field that deals with the construction of DDSs, also known as vectors, capable of carrying therapeutics to their sites of action. The action of DDSs should therefore improve the pharmacodynamic and pharmacokinetic profiles of therapeutics (2). Modern DDSs are descendants from Ehrlich's concept of a "magic bullet" (1, 8). Ideally these "magic bullets" should promote drug efficacy and reduce the emergence of toxic effects, by protecting and retaining drug integrity *in vivo*, and by constraining its release into the site of therapeutic action. Such a system must be biocompatible/ biodegradable to be stably carried throughout the body, and be able to reach the site of drug release (3). However, any drug or delivery system is subjected to harsh conditions and must override imposing frontiers during their journey through the bloodstream and other tissues. For instance, epithelia and endothelia must be crossed by transcytosis, and recognition by many of the proteins and immune factors circulating in blood must be avoided (4, 6, 9). In addition, selectivity and appropriate drug release are very difficult to translate from the research bench into *in vivo* environments and the transposition of the cell membrane is still an imposing obstacle to delivery, even for vectors of latest generation (5, 9). DDSs can be classified into two main groups (5), namely viral vectors (10), which are employed in the context of gene therapy, and non-viral

vectors (11), which can be tailored to transport a wider range of cargoes such as classical drugs and macromolecules.

1.1.1. Viral vectors

While theoretically feasible, effective gene therapy practices have been quite difficult to achieve as nucleotides are very prone to enzyme degradation (DNAses and RNAses). Furthermore these molecules cannot penetrate cells due to high abundance of anionic charges (5). In this regard, viruses were the first carriers employed for gene therapy because viruses have developed the intrinsic ability to protect, carry and deliver nucleic acids through evolution. These vectors may belong to several virus families, for example Adenoviridae, Herpesviridae, or Retroviridae, and each one possesses certain advantages and drawbacks (10, 12). In general, viral vectors have the advantage to transfect cells with efficiency as they naturally bypass cytoplasmic and nuclear membranes using cellular internalization pathways. Viral vectors can also be genetically engineered to increase selectivity. However, there are two major drawbacks associated with the use of viral vectors: I) limited loading capacity which constrains the amount of genetic material that can be carried, but more importantly II) the propensity for eliciting immune responses and the risk for induction of mutagenesis, which may not only prevent re-administration, but can seriously compromise the patient's health. Indeed, a past clinical trial with an adenoviral vector has resulted in the death of a patient, and two cases of the onset of leukemia in patients treated with gammaretrovirus were also reported in another trial (5, 10).

1.1.2. Non-viral vectors

Contrary to viral vectors, non-viral carriers may be devised to transport every type of pharmaceutical load, from nucleic acids, to proteins/peptides, and classical drugs (5). These vectors are constructed from an immense variety of structures and materials (2). Non-viral vectors can be assembled from organic molecules (e.g. liposomes, polymers, carbon nanostructures or peptides), or from inorganic colloids synthesized from such materials as iron, gold or cadmium selenide. Virus-like particles have also been constructed from chemical engineering of the natural scaffold provided by viral capsids (13).

Due to their fabrication at the nanoscale, these DDSs are collectively known as nanoparticles (NPs). NP development is usually divided into three generations. The first generation was constituted by simple colloids without surface modifications to achieve passive delivery. The

second generation includes the first targeted nanocarriers, functionalized with ligands for specific biological receptors (e.g. antibodies) in attempts to increase selectivity. The current and third generation is comprised by complex NP systems tuned to bypass physiological obstacles, to increase active targeting capacity, and to ensure some degree of temporal control of drug release through physical and chemical stimuli (2, 11). However, most of this technology must be further optimized before reaching the clinical setting, as it still exhibits considerable drawbacks. These include long term biocompatibility concerns, poor retention of colloidal stability *in vivo*, and the need for fine tuning of targeting efficiency and drug release capability (1, 9, 11).

Intracellular delivery is another aspect of NP systems that must be addressed. In fact, entering through the cell membrane, escaping endosomal compartments before enzymatic degradation, and distributing the cargo to a particular organelle are still some of the most challenging aspects of drug delivery (14). In this regard, cell uptake leading to internalization and intracellular traffic of biomolecules can occur by a diverse set of endocytosis mechanisms, divided into pathways which either proceed by non-specific entrapping of molecules on the cell surface, or which are mediated by cell-receptors. This receptor-mediated uptake is especially important for last-generation DDSs because it may be exploited for selective internalization of NPs into particular cell types. Active targeting thus involves coupling targeting moieties to NPs which are specifically recognized by particular cell-surface receptors (2). However, NP cell-targeting strategies are still quite inefficient (6, 14). Furthermore the uptake of non-viral DDSs is poorly characterized, as their internalization depends on diverse properties such as carrier size, charge, concentration or cell type (2, 6). Following cell entry by endocytosis, DDSs must escape from the endocytic vesicles. This is due to the fact that as endosomes mature, and eventually fuse with lysosomes, a subsequent drop in vesicle pH activates catabolic enzymes responsible for the degradation of molecules entrapped in the endocytic vesicles. Development of DDS escape strategies has been reported in the literature, such as pH-sensitive or fusogenic peptides (15), the construction of fusogenic lipoplexes (6), and the use of dynamic polyconjugates or proton-sponge polymers (9). Nonetheless great amounts of research must be conducted so these strategies can be properly adjusted to every NP system in order to ensure sufficient activity without eliciting cytotoxicity (6, 9, 15).

After three decades of development, few NP delivery systems are reaching the market, but more are currently under preclinical or clinical trials. Nonetheless many limitations still plague the majority of DDSs reported in the literature, which must be solved in order to increase the flux of novel developments to clinical settings (11). In this context, cell-penetrating peptides (CPPs) are an important and diverse category of biomolecules that can transport drugs through biological

membranes with high efficiency (16). CPPs have been extensively employed in the construction of third-generation NP delivery systems, as they can be chemically coupled to improve the efficiencies of uptake or to promote endosomal escape (17-20). In fact, these peptides have considerable advantages relative to other DDSs, such as being less prone to elicit immunological phenomena and not depending on receptor-mediated endocytosis as a principal pathway for uptake. Furthermore, some CPPs have demonstrated the potential to directly translocate cytoplasmic membranes independently of endocytosis, which is a rare trait in biological macromolecules (17). CPPs also present a vast repertoire of vector possibilities due to an immense variety of conceivable amino acid sequences which might be derived from natural or synthetic sources. These peptides can be extensively optimized by tailoring their sequences through mutation of selected amino acid residues (e.g. amino acid substitutions, shuffling or truncation of sequences). Several mutant variants of each CPP can thus be created, further expanding the variety of conceivable designs.

Figure 1.1 graphically depicts how CPP-based technology is deeply connected to the field of drug delivery (19, 20).

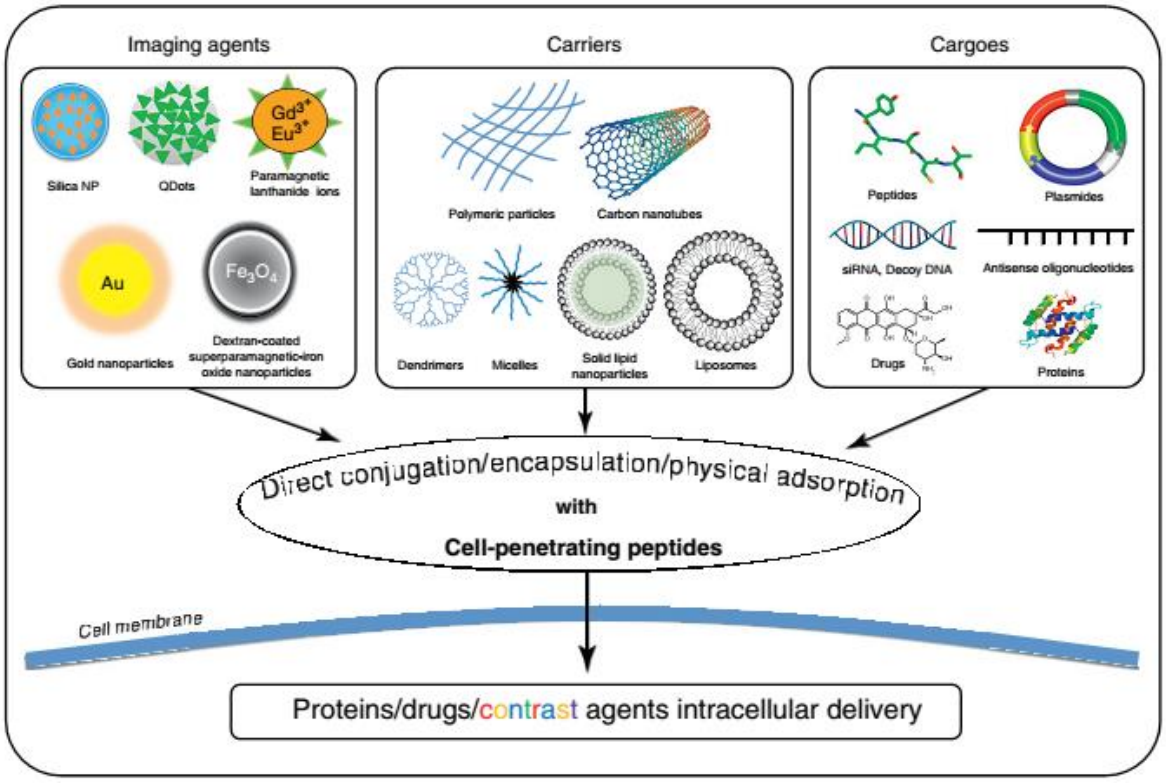


Figure 1.1 – CPPs in Drug Delivery – Representation of several biomolecular therapies and available DDSs. CPPs can be developed as DDSs or coupled to nanoparticle vectors in order to circumvent biological membranes. Adapted from (17).

In fact, it can be said that CPP history follows a parallel with that of biomedical nanotechnology, in a sense that only now, after decades of intense research and characterization by academia, have they been able to start realizing their therapeutic potential (20, 21).

For their relevance as core subjects in this work, the following chapter will address these molecular entities in detail, presenting their mechanisms of action, cargo loading strategies, and therapeutic successes and pitfalls.

1.2. Cell-penetrating peptides

Due to their remarkable properties, peptide-based drugs and drug candidates have been steadily increasing in the past decade, with more than 100 having reached the market and a few having already surpassed the global sales threshold of a 1000 million US$ (22, 23).

In this context of investment in peptide therapeutics, the acronym "CPP" refers to the most accepted nomenclature for the subset of peptides capable of interacting with, inserting into, and ultimately transposing lipid membranes independently of a need for chiral recognition by cell-surface receptors. In general, these molecules display net positive charge and are comprised by amino acid sequences between 5 and 30 residues in length, obtained from naturally occurring regions in proteins or designed by *in silico* methods. As a particular subset of leads for the development of vectors that can transport cargo molecules to the interior of cells, CPPs are already key players in the field of innovative DDSs (19).

1.2.1. History

CPP history, illustrated in Figure 1.2, had its genesis 25 years ago with the observation by Frankel and Pabo (24) that the HIV-1 transcription-transactivating protein (TAT) could enter cells and translocate into the nucleus. Three years later, it was demonstrated that the *Drosophila* Antennapedia homeodomain could be internalized by neuronal cells, leading to the discovery, in 1994, of the first CPP, penetratin (25). Afterwards, in 1997, Lebleu *et al.* (26) identified a peptide (Tat), containing the minimum functional sequence from HIV-1 TAT required for cellular uptake. The first proofs-of-concept for *in vivo* application of CPPs were reported using peptides and proteins in 1999 (27).

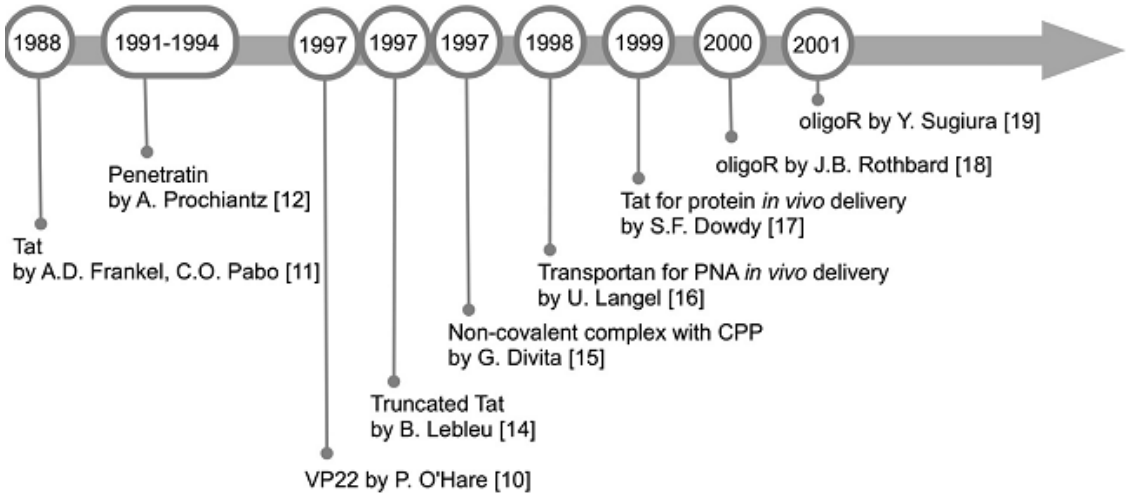


Figure 1.2 – CPP Timeline – Main events in early CPP research. Adapted from (28).

Langel *et al.* (29) introduced the "cell-penetrating" nomenclature with their work on transportan, the first chimeric peptide carrier. Several other terms have existed to refer to CPPs, such as "protein transduction regions" and "membrane translocating peptides", but those terminologies have been abandoned (30). Another cornerstone of CPP development, the strategy of using non-covalent binding between the molecular cargo and CPPs, dates back from 1997. It was first based on nucleic acid delivery mediated by MPG (a short peptide consisting of hydrophilic and hydrophobic regions), and later on the primary amphipathic peptide pep-1 to deliver peptides and proteins (16, 31, 32). A final relevant event in CPP history was the demonstration by Wender and Futaki *et al.* (33, 34) that octa-arginine sequences were sufficient for eliciting cellular and *in vivo* peptide uptake, originating R_8 and the family of poli-arginines, which are currently among the most extensively studied CPPs (30).

Ever since the discovery of penetratin and Tat, researchers have been steadily increasing the pool of known CPPs, originated from natural, synthetic or chimeric sources (28). Some representative examples are listed in Table 1.1. Furthermore, the CPPsite webpage (http://crdd.osdd.net/raghava/cppsite/) (35) constitutes a major curated database for CPPs comprising more than 800 peptide entries collected from literature.

Table 1.1 – Nomenclature, origin, sequence and cargo type of representative CPPs. Adapted from (36)

Peptides	Origin	Sequences	Cargo types
Peptides deriving from protein transduction regions			
Tat	HIV-Tat protein	PGRKKRRQRRPPQ	Protein/peptide/siRNA/ liposome/nanoparticle
Penetratin	Homeodomain	RQIKIWFQNRRMKWKK	Peptide/siRNA/liposome
Transportan	Galanin-mastoparan	GWTLNSAGYLLGKINLKALAALAKKIL	Protein/PNA/siRNA
VP-22	HSV-1 structural protein	DAATATRGRSAASRPTERPRAPAR-SASRPRRPVD	Protein
Amphipathic peptides			
MPG	HIV Gp41-SV40 NLS	GALFLGFLGAAGSTMGAWSQPKKKRKV	siRNA/ODN/plasmid
Pep-1	Trp-rich motif-SV40 NLS	KETWWETWWTEWSQPKKKRKV	Protein/peptide
MAP	Chimeric	KALAKALAKALA	Small molecule/plasmid
SAP	Proline-rich motif	VRLPPPVRLPPPVRLPPP	protein/peptide
PPTG1	Chimeric	GLFRALLRLLRSLWRLLLRA	Plasmid
Cationic peptides			
Oligoarginine	Chimeric	Agr8 or Arg9	Protein/peptide/siRNA/ODN
hCT (9–32)	Human calcitonin	LGTYTQDFNKTFPQTAIGVGAP	Protein/plasmid DNA
SynB	Protegrin	RGGRLSYSRRRFSTSTGR	Doxorubicin
Pvec	Murine VE-cadherin	LLIILRRRIRKQAHAHSK	Protein/peptide

1.2.2. Biochemical requirements for CPPs

The exact biochemical properties that confer cell-penetrating activity to a peptide have not yet been identified. This is due to a great variety in the amino acid sequences of CPPs and the fact that they are divided into families which do not share sequence similarity (30, 37). However, general biochemical profiles have been found to correlate with CPP activity, such as particular configurations in charge, hydrophobic amino acid residue content, and a propensity for folding into secondary structures in membrane environments.

In terms of the first biochemical property, net positive charge at physiological pH is the most recognizable feature of these peptides. Although there are rare exceptions (e.g. the anionic peptide SAPE (38)), the vast majority of CPP sequences are enriched in basic amino acid residues,

namely arginines and lysines, which are determinant for their activity. For example, Tat is highly cationic with 6 arginine and 2 lysine residues, and it was found that substituting any of its basic residues by a neutral alanine would markedly reduce uptake, while substitution of uncharged residues would have no effect on Tat activity (39). Furthermore, it has been found that arginines potentiate internalization relative to lysines, although often at a cost of an increase in toxicity (39, 40). Arginine is the most basic of all amino acids, because its side chain ends with a guanidinium group which allows multiple hydrogen bonding with anionic and polar molecules. Thus, arginine can form bi-dentate hydrogen bonds with phosphates moieties of more than one lipid headgroup. On the other hand, lysine can only form monodentate hydrogen bonds with molecules and therefore can only interact with a single lipid. As such, arginine rich peptides can bond with more zwitterionic and anionic lipids than their lysine rich counterparts (39, 40). The magnitude of charge is equally linked to internalization efficiency, and it has been demonstrated that in some peptides a minimum of 8 positive charges was required for efficient translocation (33).

Hydrophobic residues are also regularly present in CPP sequences, and in these cases may be critical for internalization. For instance, a study in the CPP pVEC demonstrated that single amino acid substitutions in a terminal hydrophobic patch decreased cellular uptake of the peptide (41). Moreover, other studies showed that the uptake of penetratin was abolished by substitution of tryptophan residues by a phenylalanine (7) or that the uptake of peptide R_7 was enhanced by the addition of a C-terminal tryptophan (40).

The relative importance of secondary structure in membrane insertion is still under debate. Structure polymorphism has been observed for peptides such as penetratin, reported to assume α-helical and β-sheet conformations under different conditions (e.g. different model vesicle compositions, salt concentrations, or peptide to lipid ratios (P:L)) (40, 42-44). Conversely, studies in Circular Dichroism (45), Nuclear Magnetic Resonance spectroscopy (46) and Molecular Dynamics simulations(47) have been concordant in finding that the peptide BP100 remains unstructured in aqueous environment, but promptly adopts a α-helical conformation in different lipid membranes (48). Other peptides, such as Tat and R_9, are proposed to be inherently disordered (42). Therefore structural flexibility and polymorphism may be important for amphipathic peptides (49), without being obligatory in cationic CPPs (50).

Altogether, these general considerations suggest that CPP uptake does not depend on a single parameter, but on the conjugation of biochemical conditions that render internalization energetically favorable (51), especially an equilibrium between charge and hydrophobicity.

1.2.3. Classification

Because of the high degree of heterogeneity in CPPs, conceiving straightforward classification methods has proven to be hard. Several approaches can be found in the literature, namely classification through origin, cargo-loading strategy or physical-chemical characteristics of the sequences.

In terms of origin, CPPs are classified as protein-derived, when originated from regions of protein amino acid sequences; chimeric, when resulting from the conjugation of two previously existing amino acid motifs; or synthetic, when constructed *de novo (52)*.

Relative to linkage to the therapeutic agent, CPP-cargo systems can be divided into the classes of covalent bonding and electrostatic affinity (53). In the first case, CPPs are attached to their cargo via covalent conjugation through cross-linking chemistry or through fusion of CPP tags in cloned proteins. In contrast, electrostatic affinity involves non-covalent formation of complexes due to attraction between drugs and CPPs of symmetric charges.

Concerning the physical-chemical classification criteria, Francesca Milletti (7) has recently proposed a comprehensive and systematic summary of so-far described CPPs. A representative set of 100 well-characterized CPPs, the majority (83%) of which had positive net charge, was divided into the cationic, amphipathic and hydrophobic categories (7, 17, 52). Amphipathic peptides comprised the largest of these classes, accounting for 44% of samples. Conversely, hydrophobic sequences were the rarest (15%). According to Milletti, a CPP is considered of the cationic group if it contains a stretch of positive charges that is essential for uptake, and if its secondary conformation does not lead to formation of an amphipathic structure. This class included Tat, poli-arginines and nuclear location signal (NLS) peptides. The category of amphipathic CPPs includes peptides with opposing hydrophobic and hydrophilic regions, and is subdivided into primary (amphipatic through sequence) or secondary (amphipatic through spatial conformation) α-helical, β-sheet and proline-rich peptides, according to the secondary structure that is thought to mediate their uptake. Finally, hydrophobic CPPs include peptides mainly composed by apolar amino acid residues, either having very low net charge (less than 20% of sequence) or a hydrophobic motif determined to be crucial for uptake.

1.2.4. Mechanisms of cellular internalization

The mechanisms by which CPPs enter cells are still far from resolved. Biophysical assays (43) and computational simulations (54) have shed some light on this issue at the molecular scale,

by revealing that these peptides can exert profound effects on lipid surfaces. Therefore, it has been proposed that the initial steps of CPP uptake involve electrostatic binding between basic amino acid residues and negatively charged proteoglycans or phospholipid moieties on the cell surface (39). This is followed by peptide insertion into the aliphatic regions of membrane leaflets. In many CPPs this insertion step is thought to be related to the influence of hydrophobic residues, or to conformational shifts which lead the peptides to acquire an amphipathic structure (30, 39). Proceeding stages of the molecular mechanism of cellular entry of CPPs are equally unclear.

The first studies conducted to ascertain the biological entryway for Tat, penetratin and R_9 used microscopy observations, and fluorescently activated cell sorting, on cells incubated with fluorophore-labeled peptides (40). These studies suggested that CPP *in vitro* uptake could take place despite treatment of cells with endocytosis-blocking drugs. Such findings lead to the conclusion that these peptides could directly translocate membranes through an energy-independent, non-endocytic process (36, 39). However, in 2003, Richard *et al.* (55) provided evidence that formaldehyde fixation prior to microscopic analysis drastically changed the intracellular distribution of CPPs. This finding, along with the fact that flow cytometry cannot distinguish between membrane-bound and internalized fluorophores, demanded a reevaluation of the mechanisms that had been previously described (7, 54, 56). Furthermore, it is now recognized that peptide-membrane interactions, cellular uptake and cytosolic distribution, are influenced by such factors as the concentration of the peptides, lipids and salts in the solvent, medium pH, lipid membrane composition or cell type (57, 58) and that cargo molecules, including fluorescent probes, also substantially impact CPP internalization routes (59, 60). Thus each CPP/cargo DDS has its own particular behavior and must be studied independently. Despite this drawback in unearthing the biological pathways responsible for CPP internalization, the vast number of publications in the literature concerning the biological and biophysical characterization of CPPs has resulted in the formulation of general models for the mechanisms of uptake. Figure 1.3 illustrates some of these models. Thus, CPP cellular uptake is thought to proceed through endocytosis or through direct translocation of cell membranes.

The left panel of Figure 1.3 comprises several endocytic mechanisms which have been implied in CPP internalization. Endocytic mechanisms are natural processes occurring in living cells that can be triggered by electrostatic interactions with proteoglycans in the extracellular matrix, by direct interaction with the plasma membrane or through specific binding to cell-surface receptors (40). The influence of CPP interactions with proteoglycans and the plasma membrane as triggers for endocytosis remains unresolved (61-63).

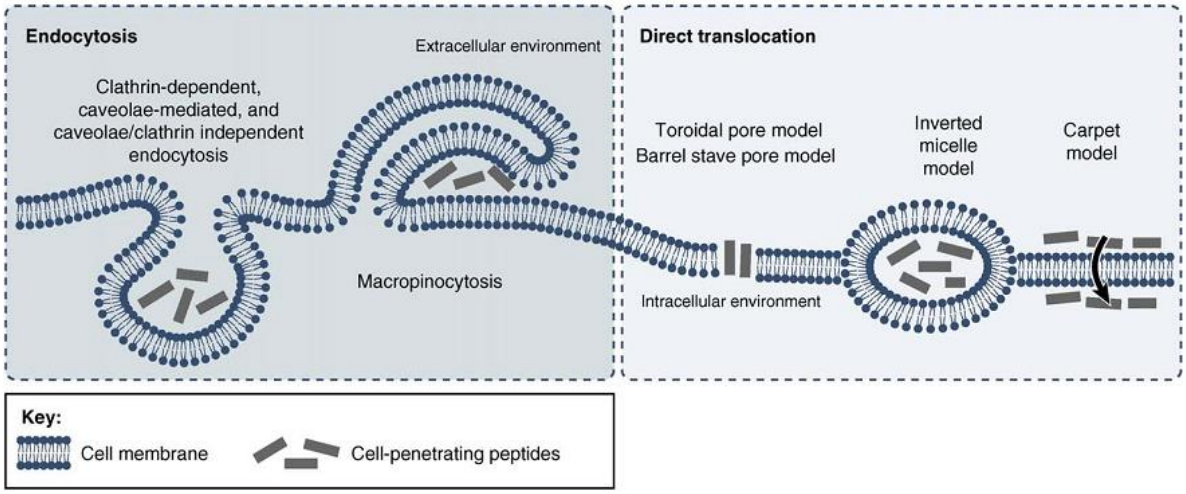


Figure 1.3 - CPP cellular entry routes and models of uptake - CPP internalization can be divided into energy-dependent and energy-independent pathways. The first type comprises clathrin-dependent, caveolae-mediated and caveolae/clathrin independent endocytosis, and macropinocytosis. Direct translocation may proceed by the torroidal pore model, the barrel-stave model, the carpet model, or via inverted micelles. Adapted from (64).

On the other hand, receptor recognition of target molecules is dependent on their chirality, and both L- and D- stereoisomers of a CPP can be equally internalized under the same testing conditions (52), i.e it is clear that receptor-mediated endocytosis does not contribute for CPP internalization. As illustrated in Figure 1.3, all of the currently recognized pinocytic routes have been implicated in CPP uptake (40). For instance, macropinocytosis involves an actin-driven invagination of the plasma membrane which forms large irregular vesicles, and several studies have supported its involvement in the uptake of poly-arginines, penetratin and Tat (40). Clathrin-mediated endocytosis, formed by energy-dependent assembly and pinch-off of clathrin-coated pits from the plasma membrane, has also been connected to Tat and penetratin (39). Caveolin-dependent endocytosis, on the other hand, is lipid raft-mediated and depends on hydrophobic membrane cholesterol-rich microregions, the caveolae, which bud off the membrane to form caveosomes (40). In this case, co-localization studies of fluorescently-labeled Tat with cholera toxins, which are known to internalize via caveolae-mediated endocytosis, supported the uptake of Tat through this entryway. However, contrary results were reported by authors who found that treatment with inhibitors for caveolae-mediated uptake was unable to abolish Tat internalization (39). Examples like these reflect the often contradictory results that have been obtained from different CPP uptake studies. In fact, no single endocytosis pathway could yet be resolved as a predominant contributor

for CPP internalization, and it is now though that several different endocytic mechanisms may act in concert to promote the uptake of each CPP or CPP/cargo system.

The rightward panel of Figure 1.3 concerns non-endocytic internalization. Although direct translocation had been related to artifacts of cell fixation in early studies of CPP uptake, more recent assays on live cells have re-established its involvement in the internalization of some CPPs (39). In this regard, direct translocation involves peptide action in transiently destabilizing and bypassing cell membranes independently of ATP consumption or endocytosis. Figure 1.3 displays four pathways which have been proposed for direct translocation of CPPs, namely the torroidal pore, barrel stave and carpet models, as well as the inverted micelle model. According to the barrel-stave model, translocation happens when a number of peptides form a membrane channel by self-assembly into a barrel-like ring around an aqueous pore. The toroidal pore model is similar as it involves the formation of another type of pore, triggered by peptide induction of high curvature folds in the bilayer. On the other hand, the carpet model results from an accumulation of peptide molecules electrostatically bound to the membrane surface so that when a P:L threshold is reached, the barrier is locally destabilized without the formation of pores (65). Finally, the inverted micelle model states that disturbances in lipid bilayers lead to the formation of inverted micelles that trap the CPPs in their hydrophilic core until further destabilization, effectively releasing the molecules into the cytosol (64, 66). Additional models for peptide transposition mechanisms, e.g. driven by membrane potential differences, can also be found in the literature (67). Nonetheless, it should be noted that, although informative, such models are limited approximations, and thus may not fully describe the complex molecular interactions that take place during CPP activity (68).

As a consequence of the former paragraphs, it should be realized that several kinds of endocytosis and translocation mechanisms may act in concert to promote the internalization of each CPP/cargo system. Indeed, this notion is gaining increasing support and acceptance in CPP-related literature (20, 39).

1.2.5. CPP-mediated gene therapy

CPP delivery strategies have gained considerable popularity in the past decade, and more than 1000 applications have already been attempted both *in vitro* and *in vivo* (54). In fact, since the first *in vivo* proofs-of-concept, CPPs have been used in therapies to target multiple disorders including asthma, ischemia, diabetes, or inflammation (54, 56). Therapies for cancer have also been intensively studied, and CPPs are equally investigated for treatment of brain related illnesses

because of their ability to cross the blood-brain barrier (69, 70). In this regard, the application of CPPs as gene and oligonucleotide therapy vectors is a particularly promising and widely researched area (28).

Gene therapy is a powerful form of treatment by which cells are supplied with exogenous genetic material that complement, substitute, or suppress functions linked to disorders (5). Cancers (e.g. melanomas, carcinomas, leukaemia), genetic (e.g. cystic dibrosis, X-linked severe combined immunodefficiency disorder), auto-immune (e.g. diabetes type 1), neurodegenerative (Parkinson's disease, multiple sclerosis) and cardiovascular disorders and viral infections are just some of the highly problematic illnesses that may be overcome by a continued investment in this type of therapy (5, 10, 71). Diverse approaches to gene therapy have been developed in the past twenty years (72), namely the use of plasmids containing genes of therapeutic interest, or splicing correction and silencing of deleterious genes via antisense oligonucleotide technology. These strategies may be achieved using single strand DNA (ssDNA), small interfering RNAs (siRNA) or microRNAs, or by steric blocking DNA mimics, such as phosphorodiamidate morpholino oligomers (PMO) and peptide nucleic acids (PNA) (72-75). Extensive examples of CPP conjugation to all of these molecules for gene delivery, antisense knockdown or mRNA splice correction can be found in the literature (36).

For instance, CPP -mediated delivery of plasmid DNA has been achieved in cultured cells, chiefly through non -covalent approaches that have included CPPs such as MPG, Tat and poly-arginines (36). *In vivo* results have also been reported. For instance, the intravenous injection of a luciferase expression plasmid conjugated with the peptide PPTG led to significant gene expression in the mouse lung (76). In another study, a recombinant form of Tat was mixed with a therapeutic plasmid encoding α-galactosidase A and injected into muscles of knockout mice models for a lysosomal storage disease. In this latter case the skeletal muscle expression of galactosidase was significantly enhanced in comparison to injection of DNA alone (58).

Likewise, CPP strategies have been developed for the administration of siRNA *in vitro* and *in vivo*. For example, MPG has been reported to enhance siRNA delivery in several cell lines, but has also been applied for *in vivo* delivery of OCT-4 targeting siRNA into mouse blastocysts (77). Furthermore, Tat conjugated with an RNA-binding motif was reported to block epidermal growth factor *in vivo* (78), and the fusion of a small peptide from rabies virus glycoprotein to R_9 was used to deliver siRNA to the CNS of mice (79).

Finally, the conjugation of oligonucleotide mimics to CPPs has also been extensively assayed in peptides such as transportan Tat, penetratin and oligo-arginines (36). *In vivo* results have

also been reported. Prominent examples include several synthetic and chimeric peptides which have been tested in the delivery of therapeutic PMOs to animals models of Duchenne muscular dystrophy (80).

These are only a few instances of a large amount of proofs-of-concept related to the viability of CPP-mediated gene delivery. Although encouraging results have been attained, CPP research still has an important role in the translation of these applications from the work bench to the clinical setting (21).

1.2.6. Current drawbacks of CPPs

Despite demonstrating remarkable properties and constituting promising vectors, CPPs encounter obstacles to clinical translation. Lack of selectivity, and concerns about toxicity, are perhaps obvious, but problems like reduced *in vivo* efficacy, poor bioavailability, instability, and short time-span, must also be considered (7). A number of studies have assessed the toxicity of CPP *in vitro*, while *in vivo* information from animal assays remains more scarce (58, 81-84). In general CPP have demonstrated low levels of cytotoxicity (85, 86). The kinetics and biodistribution of some CPPs have also been studied *in vivo* establishing that these therapeutics can potentially be safe and target most tissues in the body (64, 87-89). However, downsides such as rapid blood clearance and preferable distribution to the liver, kidneys, spleen, bowels and lungs, have also been reported (90, 91). In this regard, phagocytosis by the reticuloendothelial system, proteolysis by intestinal enzymes, as well as interactions with blood components can severely decrease the circulation of CPPs.

However several strategies are being investigated to counter CPP drawbacks, such as substitution of biological L- amino acids for protease resistant R- stereoisomers, chemical optimization of structure, or covalent linkage to shielding molecules, such as poly-ethylene glycol (21, 64). Clearly CPPs have demonstrated great therapeutic potential but there is still the need and ample opportunities to expand the known pool of peptide leads and to continue studying the biochemical properties which are determinant for their activity.

1.3. Structural viral proteins and CPPs

It was previously mentioned that the discovery of TAT protein's natural translocation ability is often held at the beginning of CPP research. Additionally the Tat peptide was one of the first CPPs to be discovered and one of the most extensively studied peptide DDSs (40). However,

more examples exist of CPPs derived from viruses, concretely from structural components of viral shells. Vp22 is one such example, found in a Herpes Simplex virus tegument protein capable of translocating membranes (92). Additional CPPs have been originated from the glycoprotein H of the same virus (93), or from the k8 protein of Herpesvirus 8 (94). Other sources of CPPs from structural viral proteins have included the classical swine fever virus' envelope protein (95), the human respiratory syncytial virus surface glycoprotein G (94), the hepatitis-B surface glycoprotein (96) and a N-terminal domain of its X protein (97), or the capsid units of the flock house virus (98), the brome mosaic virus (99), and the alphavirus (94). Furthermore the extensively studied chimeric CPPs pep-1 and MPG comprise a NLS from the Large T antigen from simian virus 40 (16), and the hydrophobic motif of MPG is derived from the fusion sequence of the HIV protein gp41. Nonetheless, despite these well-established precedents, the percentage of CPPs derived from viruses is still very limited when compared to the total number of currently known CPP sequences, as indicated in Figure 1.4. This suggests that structural viral proteins are an underexplored source of therapeutic peptides.

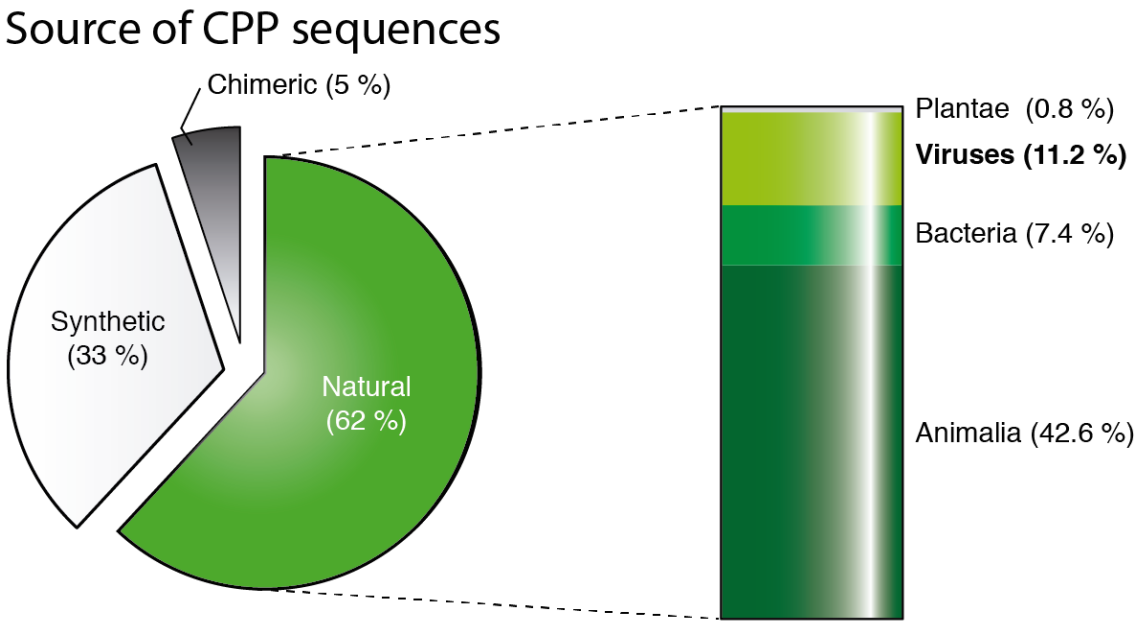


Figure 1.4 - Sources of CPP Sequences – Each percentage was calculated based on the amount of hits on CPPsite (35).

In fact, viruses are fascinating, complex and very diverse molecular entities that have evolved, through aeons of selective pressure, a wide array of strategies to protect and deliver

genomes to their replication sites in the interior of obligatory cellular hosts (100). Moreover, viruses persist throughout all of life's regions and exhibit a large number of taxa with differing sizes and morphologies. Simple viruses can possess a single genome molecule contained within a protective structure, the capsid or nucleocapsid. These structures may be organized in a variety of arrangements, such as helical or icosahedral patterns (101). More complex virions have additional elements, including non-structural proteins with diverse roles in the viral cycle (102) (e.g. genome replication, modulation of host gene expression), and the envelope (103), a lipid bilayer surrounding the viral core in association with a protein shell. A distinction between enveloped and naked viruses is important because the strategies to transpose cell barriers employed by species within each group may be quite distinct (104), as illustrated in Figure 1.5.

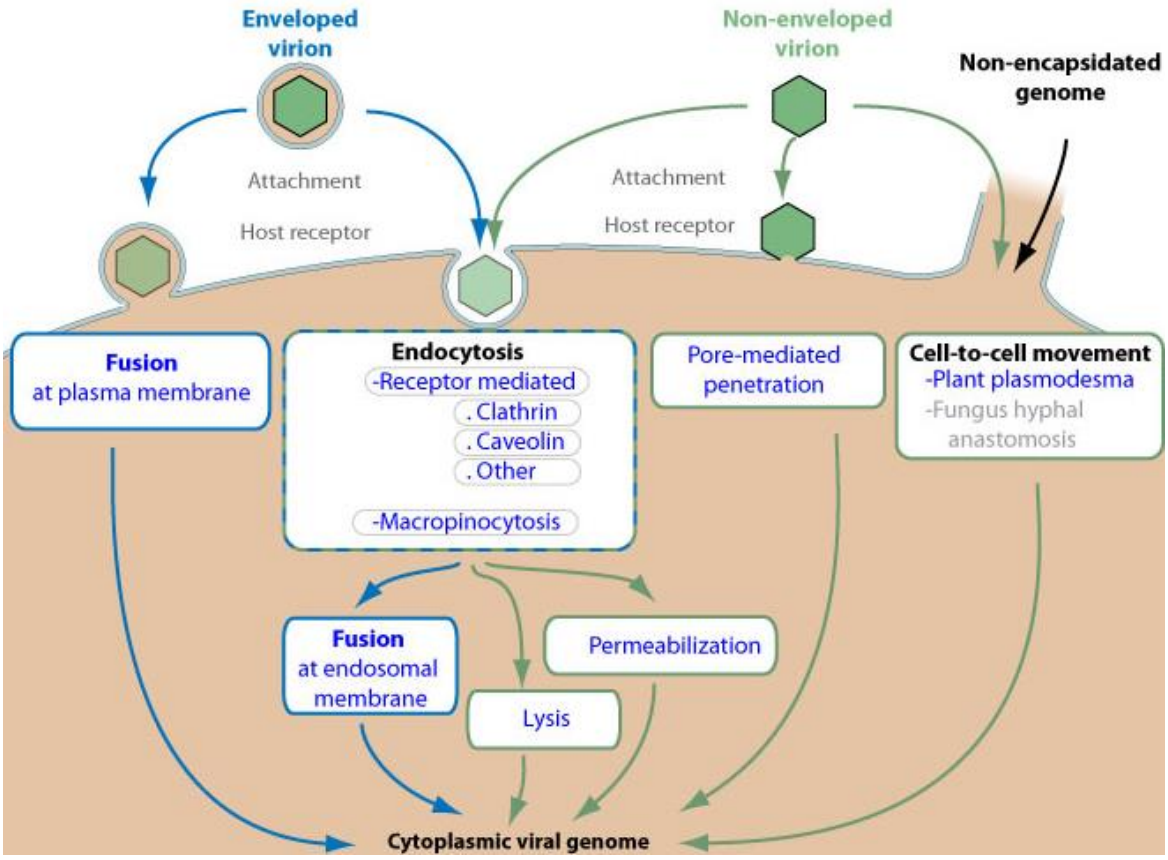


Figure 1.5 - Diversity in viral cellular entry mechanisms – Structural proteins from capsid, envelope or membrane elements mediate the uptake of virions into host cells through a variety of strategies. Adapted from Expasy ViralZone (105).

As the figure indicates, some viruses gain access to the cytosol through the cytoplasmic membrane (106), while others are first internalized and transported through endocytic vesicles (107). In addition, viruses have also been shown to exploit the proximity between neighboring cells, spreading from the interior of a host directly through the membrane to the cytosol of contacting targets (100, 107). Therefore, a fundamental consequence of viral uptake is that virions must interact with lipid bilayers to reach their replication sites. In this regard, several mechanisms have already been characterized, including the induction of membrane lysis or of fusion events, and the formation of pore structures through which genomes can escape (108).

Due to the propensity that structural proteins from viruses display towards lipid membranes, some authors have recognized the similarities between selected regions of these proteins and a broad group of peptides, collectively known as membrane-active peptides (MAPs) (109-111). These peptides share the intrinsic property of exerting biological effects (antimicrobial (45), antiviral (112), anticancer (113)) based on molecular interactions with lipid membranes. Indeed, CPPs can equally be classified as MAPs because their activity is related to an affinity for lipid structures (109, 114). Given this fact, and the data presented in Figure 1.4, it is surprising that viruses have not been more intensively studied in the field of CPPs, and in a wider perspective, of therapeutic MAPs.

Professor Castanho's lab has combined the subjects of structural proteins from viruses and MAPs in the determination of Flavivirus entrance mechanisms through the study of the dengue virus (DENV) capsid protein. DENV, an enveloped virus, has three structural proteins termed as envelope, membrane, and capsid, relating to the structures from which they are originated. There is ample evidence that this virus enters cells through endocytosis and that endossomal acidification triggers a fusion type II mechanism (115). This entry strategy consists in the formation of envelope protein trimers and their insertion into the host membrane, but consecutive stages of the process are still left up to speculation (116). Recently, Freire *et al.*, from Professor Castanho's group, have found that the C protein is implicated in Flavivirus fusion by demonstration of its inherent capacity to permeate cells. For this end, peptides were first derived from the two putative functional domains of the C protein monomer, the RNA binding domain – pepR, and the membrane binding domain-pepM. Both peptides were found to form stable electrostatic aggregates with ssDNA and to strongly favor partition into model membranes, even when conjugated with nucleic acids (111, 117). Furthermore pepR and pepM were confirmed as full-fledged CPPs capable of penetrating cells and delivering ssDNA with high efficiency (118). Uptake studies suggested that pepR internalization is dependent on an endocytic pathway, and that pepM can directly translocate cell membranes in

accordance with the inverted micelle model (118). PepR had also previously been reported as an antimicrobial peptide (AMP) capable of severely disrupting the outer surface of *Escherichia coli* (119). Furthermore, the full-length DENV C protein was also studied and shown to directly translocate cells while delivering nucleic acids *in vitro* with high efficiency (120). This work also proposed a relation between the capsid of Flavivirus and a recent concept of super charged proteins (SCP). Liu *et al.* have found that engineering natural proteins like green fluorescent protein (GFP) to express a very high density of positive charges per molecular weight (e.g. GFP wt: -8 to SCP GFP: +36 charges) can impart them with the ability to deliver functional cargo to cells, outperforming even cationic CPPs (121-124). The authors also found SCPs expressed in the human genome, and tested them as DDSs *in vitro*, effectively establishing that SCPs can occur spontaneously in nature (125). Freire *et al.* have further expanded this idea by identifying the DENV C monomer as a SCP. In fact, DENV C protein has a charge to weight ratio of +1.97 kDa/MW (120) and this value greatly surpasses the cutoff for a protein to be considered super charged, set by Liu *et al.* (of > +0.75 kDa/MW)(125). Freire *et al.* also studied the prevalence of natural SPCs in structural proteins from additional viral families (120) (categorized as elements from envelope, membrane or capsid). They found that the great majority of SCPs in the structural viral proteome corresponded to capsid elements, including capsids from all other members of the Flavivirus family. Additionally it was found that, excluding Tat, all virus-derived CPPs did not belong to a natural SCP (120).

1.4. Motivation

The work described in this document further explores the findings of Freire *et al*, that viral proteins and/or derived peptides are efficient intracellular carriers of genetic cargo. Thus, the main objective was to identify specific peptide sequences in structural proteins from viruses with the inherent property of interacting with lipid bilayers, and exploit them as delivery vectors for gene therapy. To do so, the project assumed a tripartite configuration, as described in the following paragraphs.

The first stage of the work concerned the identification of putative CPP regions (viralCPPs) in capsid, envelope or membrane proteins of viral proteomes. This was achieved by collecting a large sample of structural viral protein sequences from online databases. These were then uploaded to the CellPPD website (126), containing an algorithm for prediction of CPP regions in protein sequences. This process resulted in thousands of putative CPP sequences, or viralCPPs. These

sequences were grouped together based on pairwise similarity using the software package CLANS (127). A total of 150 viralCPPs were selected from the several clusters produced by CLANS. A final set of 14 viralCPPs was then chosen from the previous group, to be assayed *in vitro* and validated as CPPs.

As such, the second stage of the project involved the execution of a protocol for *in vitro* assessment of the 14 viralCPPs to evaluate their capacity in the delivery of ssDNA oligonucleotides into cells. Consequently, the peptides were tested as possible DDSs for gene therapy applications. This protocol consisted on mixing each viralCPP with Alexa-488 labeled ssDNA oligonucleotides to create electrostatically bound complexes, adding them to cells, and incubating the system for one hour. After washing, to remove non-internalized molecules, the fluorescence intensity of probe retained by cells was evaluated by fluorescence spectroscopy and flow cytometry. Therefore, this protocol allowed that the efficiency of viralCPP-mediated ssDNA delivery be evaluated. Confocal microscopy was employed to visualize viralCPP delivery of oligonucleotide probes. The effect of trypsin on viralCPP activity was also studied, and cytotoxicity assays were conducted to establish the influence of viralCPPs on cell viability.

Finally, in the third stage of the work, biophysical methodologies were conducted to assay biochemical properties of these peptides that could be related to their biological activity. Di-8-ANEPPS membrane dipole potential sensing experiments were made to assay the degree of interaction between each viralCPP and lipid model vesicles. This study was important because membrane affinity is linked to CPP activity. Furthermore, circular dichroism spectroscopy was employed to detect conformational shifts in viralCPPs in the presence or absence of lipids, a common biochemical trait in CPPs. A last set of experiments used dynamic light scattering spectroscopy to calculate the hydrodynamic diameter of viralCPP/ssDNA electrostatic complexes. Hydrodynamic diameters relay the apparent size of peptide/cargo conjugates which can be linked to the efficiency of delivery.

This work is of relevance in several areas of biomedical research. Indeed, new CPP leads could be identified and isolated from structural viral proteins, a previously underexplored source. In this regard, the project marked a “return to origins” in the context of CPP research due to the fact that its cornerstone - HIV’s TAT– was a viral protein. Furthermore, proof-of-concept was provided for viralCPP development as DDSs for gene therapy, a state-of-the-art application in medicine.

2. Technical overview

Experimental procedures conducted in this work concerned: I) the use computational tools to identify novel CPP sequences, II) an *in cellulo* approach to study CPPs combining flow cytometry (FC) and fluorescence spectroscopy (FS), III) biophysical studies to assay interactions between CPPs and lipid membranes or oligonucleotides. This section provides a brief theoretical background for the techniques that were employed during the project.

2.1. Computational tools

2.1.1. Expasy ViralZone

ViralZone (http://viralzone.expasy.org/) (105) is a web resource which provides fact sheets on all known virus families, combining textbook knowledge with genomic and proteomic sequences. For each genus within these families, reference strains have well documented and annotated functional and structural data.

In this work, ViralZone was employed to easily locate a wide range of structural viral proteins with clear functional and structural annotated information. Using the Baltimore classification system as a starting point, viruses from all known families were accessed. For each virus the amino acid sequences corresponding to its structural proteins were collected from the UniProtKB (128). These structural proteins included elements from capsid, envelope or membrane, according to individual virion morphologies.

2.1.2. CellPPD

Understanding if a peptide can act as a CPP based on its amino-acid sequence is a non-trivial problem. This is due to a lack of clear homology in the hundreds of very diverse sequences recognized as CPPs, despite the fact that several biophysical studies have revealed biochemical features which are common to the majority of CPPs (129). These include properties such as positive net charge, relative degree of hydrophobicity and short length, which can be used as general rules of thumb for predicting CPP potential (130). Recently, important developments have been made in the field of CPP prediction. The onset of CPPsite (35), a database for CPP sequences, as well as a growing number of CPPs reported in the literature, has allowed the refinement of prediction algorithms. As a consequence, there are now two online tools for predicting CPP potential in amino acid sequences, CPPpred (131) and CellPPD (126).

CellPPD (http://crdd.osdd.net/raghava/cellppd/) (126) has been used in this work to identify putative CPP regions in structural proteins from viruses. The algorithm implemented in CellPPD divides proteins into sequences of a length defined by the user (e.g. sequences of 20 amino acids) and each sequence is analysed for its CPP propensity. Therefore, the algorithm predicts and relays the likelihood that each of those sequences is a CPP. This prediction is conducted by a machine learning algorithm instructed by a main dataset of 700 peptides. The algorithm accounts for factors such as total amino acid constitution, and relative position of particular amino acid residues in the sequence. The evaluation conducted by the CellPPD algorithm also attributes an estimator of CPP likelihood, the Support Vector Machine (SVM) score, to each sequence. Per standard definitions, protein sequences scoring positive SVM score values are considered CPP regions, and can be isolated as potential novel CPP sequences. Additionally, higher SVM scores are related to increased CPP likelihood while protein sequences with negative SVM score are considered non-CPPs. This algorithm has a probabilistic nature, and the authors report that it can predict CPPs with over 80% accuracy. Therefore, each positive prediction by CellPPD must be subjected to experimental validation to ascertain that it is a true positive hit, i.e. validate CPP activity.

2.2. Flow cytometry

FC is a technology that measures several physical characteristics of populations of living cells by means of spectroscopic data (132). To fulfill this function, a standard FC apparatus is equipped with a fluidics system able to transport single cells through the optical path of excitation lasers. The signals generated by the optical system are then individually recorded by a set of detectors and digitally processed. Data from FC allows the quantification of particle size and relative granularity/internal complexity through light scattering (as represented in Figure 2.1). Forward scattered light, scattered off the axis of an incident beam, is proportional to cell size. Side-scattering, which is refracted or reflected light, can be related to granularity and cellular complexity. The correlation between forward and side scattering allows the identification of different populations in a heterogeneous pool of cells (133). Fluorescence intensity can also be recorded by FC when cells or biomolecules are labeled with a fluorescent probe. In this case, different populations can be recognized based on fluorescent staining.

Data from a FC assays can assume different graphical representations, but histograms and 2D-scatter plots are the most frequent. In scatter-plots the concept of gating is especially important because thresholds can be defined to establish cell populations. Thus, a gate is a graphical/numeric

boundary that defines the range of values in which each population is included, so that positive and negative events may be distinguished and analyzed.

In the context of this work, FC was used to quantify the effect of CPPs on the delivery of ssDNA labeled with a fluorescent probe Alexa-Fluor 488 (A488). This technique was also employed to conduct toxicity evaluations through a Live/Dead assay (134).

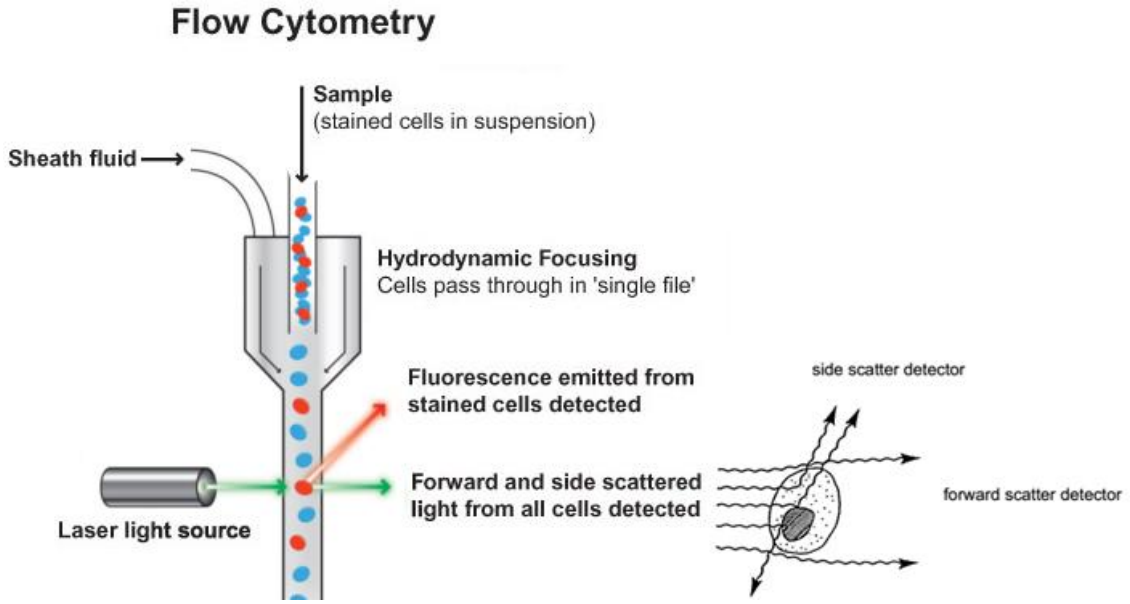


Figure 2.1 - Schematics of basic components of a flow cytometer – Samples are loaded and carried by a sheath fluid. Hydrodynamic focusing produces a single file in which particles are individually probed by a system of lasers. Fluorescence intensity and forward/side scattering of light can thus be recorded and transduced to the electronic component of the device. Adapted from (132).

2.3. Biophysical methodologies

Biophysics is an interdisciplinary field concerning the study of living systems by employing theories and techniques from physics. Thus it provides a quantitative analytical approach to many biological questions. In this context, an assortment of spectroscopic methods can be used to inquire relationships between the physical properties of molecules, such as structure, dynamics and function, in order to determine their role in cells, tissues and organs.

2.3.1. Steady-state fluorescence spectroscopy

Fluorescence results from the absorption of light by a molecule, its excitation, and the subsequent emission of light of higher wavelength after loss of energy by nonradiative processes

(135) and can be used to assay biological systems by fluorescence intensity, spectrum, lifetime or polarization measurements. It is ubiquitous in life sciences research (135, 136) and has made important contributions in the study of macromolecules. FS has been particularly seminal in the field of membrane-active peptides (137), for it allows to quantify interactions between peptides and lipids, for example through the determination of partition coefficients (43). The applications of these techniques can be divided into those which depend on the intrinsic fluorescence of aromatic amino acids residues (phenylalanine, tyrosine and tryptophan), and those which depend on the extrinsic fluorescence of molecular probes.

In this work, extrinsic fluorescence methodologies were used. In an initial set of experiments, ssDNA molecules were labeled with A488 and FC was used to collect its fluorescence intensity (A-488 spectrum is shown in Figure 2.2) and relay the efficiency of CPP-mediated ssDNA delivery *in vitro*. FS was also used in membrane dipole potential sensing experiments, described in the following paragraphs.

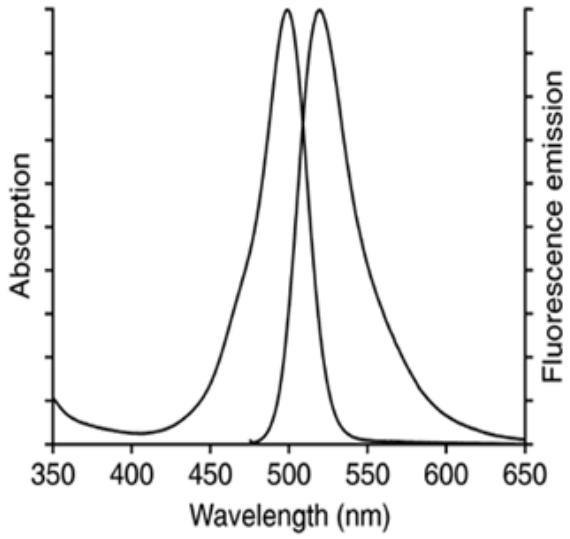


Figure 2.2- Alexa-Fluor 488 Absorption and emission Spectra

Membrane dipole potential sensing

Another application of FS in this work consisted in using the membrane dipole sensitive fluorescent probe di-8-ANEPPS to assay the effect of CPPs on lipid membrane models.

The electrical potential profile across the cytoplasmic phospholipid bilayer is comprised by the transmembrane potential ($\Delta\psi$), the surface potential (ψ_s) and the dipole potential (ψ_d) (Figure 2.3). The membrane dipole potential, in particular, is generated on the threshold region between

aqueous environments and lipid hydrocarbon non-polar chains. It results from the alignment of dipoles in water molecules hydrating the outer surfaces of biological bilayers with dipolar residues in the lipids (138, 139).

Macromolecules capable of exerting an effect on lipid bilayers can directly interfere with membrane organization or contribute with their own dipoles to the electrical potential of the membrane. These molecular interactions result in the alteration of the membrane dipole potential. Therefore, dipole potential sensing can be exploited to report interactions between macromolecules and lipid bilayers. Fast-response voltage-sensitive fluorescent dyes are typically used in dipole potential sensing assays. Among these dyes is the styrylpyridinium probe di-8-ANEPPS (4-[2-[6-(dioctylamino)-2-naphthalenyl]ethenyl]-1-(3- sulfopropyl)- pyridinium). Di-8-ANEPPS possesses a dipole-sensing chromophore, as well as two aliphatic chains that facilitate integration into biological membranes, as illustrated in Figure 2.3 (140).

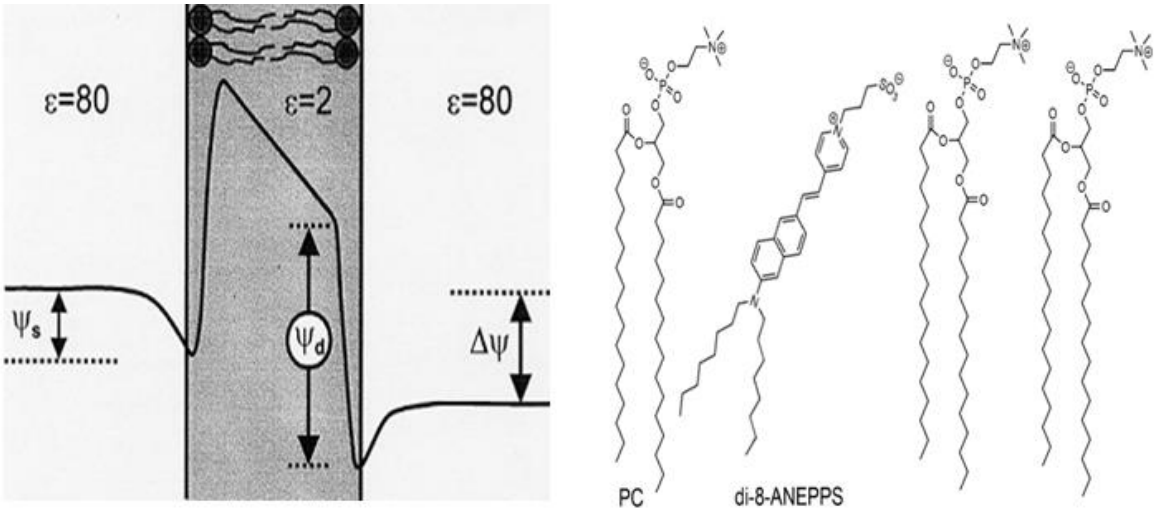


Figure 2.3 – Electrical potential across the cytoplasmic membrane (Left panel) Schematics of a di-8-ANEPPS molecule in membrane environment (Right panel). Adapted from (138).

Di-8-ANEPPS excitation spectrum shifts in response to changes in dipole potential (141). As illustrated in Figure 2.4, red shifts in the excitation spectra (to higher wavelengths) indicate drops in dipole potential, while blue shifts (to lower wavelengths) signify a gain. Differential spectra are created to report these shifts, and the underlying interactions. This is accomplished by subtraction of the excitation spectrum of di-8-ANEPPS in labeled membranes to the spectrum in membranes in the presence of interacting molecules. Both spectra are area-normalized prior to

subtraction to reflect only spectral shifts. The amplitude of the differential spectrum is correlated with the magnitude of dipole potential variations, and therefore to the extent of membrane interactions. Furthermore, the fluorescence intensity ratio (**R**) of both peaks in the differential spectrum can be used to quantify the magnitude of interaction (137). This procedure, schematized in Figure 2.4, is called a dual wavelength ratiometric assay.

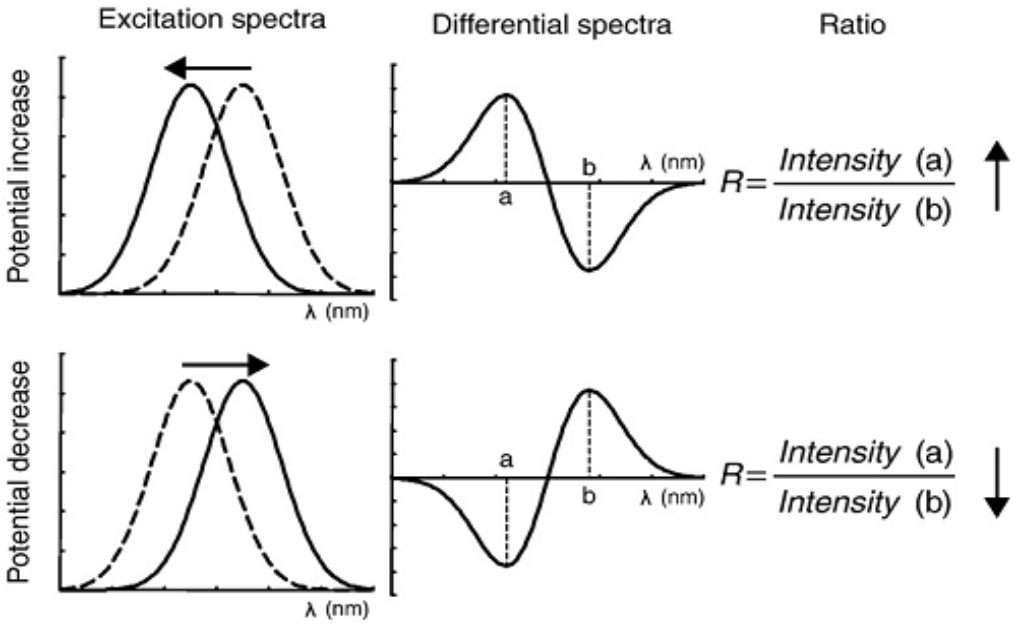


Figure 2.4 – Voltage-sensitive probe dual wavelength ratiometric measurements - Differential intensity spectra are created by subtracting the excitation spectrum of di-8-ANEPPS in model vesicles (dashed line) to the one resulting upon interaction (continuous line). Typical sinusoidal curves are obtained, indicative of either a blue or red shift. The ratio (**R**) of fluorescence intensities corresponding to maximum and minimum wavelengths (a/b) in spectra are used to quantify magnitude of interaction. Adapted from (137).

2.2.2. Circular dichroism spectroscopy

Circular Dichroism (CD) is used to determine protein and peptide structures in solution. A CD signal is originated from chiral chromophores which exist in many biomolecules. Optical activity is associated to the asymmetric environment that is generated in proteins and peptides by the adoption of a secondary structure (142).

CD spectropolarimeters measure the difference in absorbance of the left (L)- and right (R)-components of circularly polarized light ($\Delta A = A_L - A_R$). ΔA is used to estimate ellipticity θ, related to the angle between axes that describe the polarization state of light after its interaction with optically active moieties (elliptical polarization, Figure 2.5). In peptides and proteins, the peptide

and disulfide bonds and aromatic amino acids side-chains act as chromophores. Therefore, by collecting a CD spectrum in the far UV region (240–190 nm), peptide bond absorption is used to assess the formation of common structural motifs (e.g. α-helices and β-sheets) because they assume characteristic spectra, represented in Figure 2.5 (136, 142, 143). CD results are often reported as mean residue ellipticity,$[\theta] = \frac{Mr}{10 \times C \times d}\theta$, calculated from mean residue molecular weight (Mr), peptide concentration (C) and optical path of light (d).

CD was thus used in this work as a way to assess peptide secondary structure upon membrane binding.

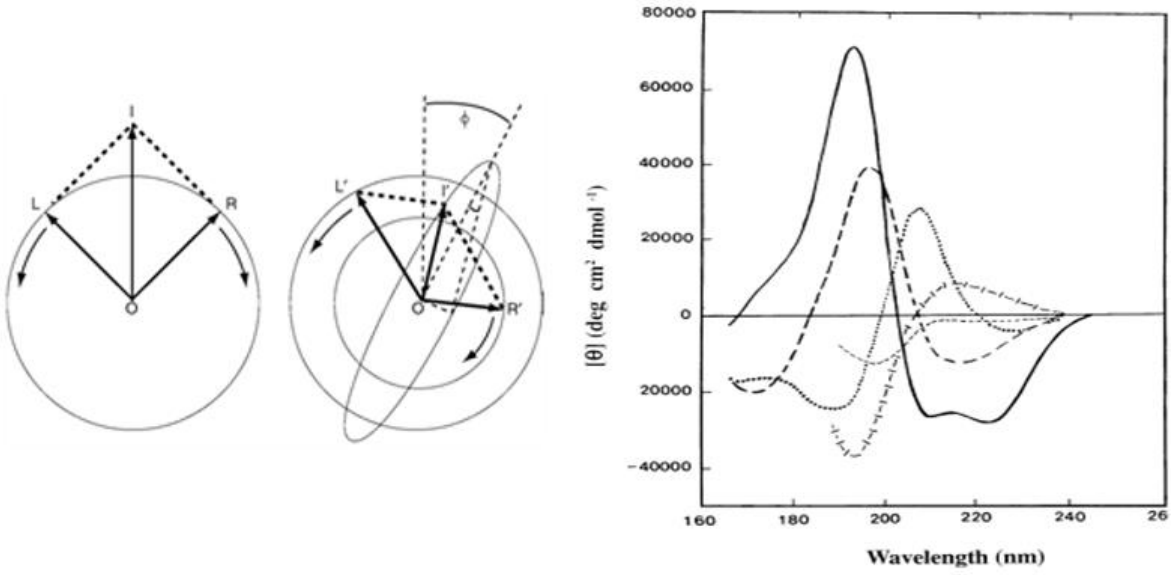


Figure 2.5 – Elliptical polarization of light (Left panel) Differential absorption of left and right components of circularly-polarized light results in elliptical polarization of the beam. Idealized CD spectra of protein secondary structures (Right panel) α-helices display large CD bands with negative ellipticity at 222 and 208 nm, and positive ellipticity at 193 nm, β-sheets exhibit a broad negative band near 218 nm and a large positive band near 195 nm, while disordered extended chains have a weak broad positive CD band near 217 nm and a large negative band near 200nm. Adapted from (142).

2.2.3. Dynamic light scattering

Dynamic light scattering (DLS) measures microsecond fluctuations in the intensity of light that is scattered while transversing an optical path through a solution of biomolecules. These fluctuations are generated by random collision of particles (Brownian motion) which is influenced by their size, medium viscosity and temperature. DLS instruments calculate an autocorrelation analysis function from these fluctuations of scattered light. This function registers the correlation in

the intensities of scattered light detected in-between increasingly longer intervals (delay times) and relates it through an exponential decay curve. The averaged diffusion coefficient (D) is relayed by the decay rate of this curve. The Hydrodynamic radius (Hr) of the scatterers in solution can then be determined by the Stokes-Einstein equation, $D = \frac{kT}{6\pi\mu Hr}$, where k stands for Boltzmann's constant, T the absolute temperature and μ the medium viscosity (136, 144). It should be noted that the Hr is the radius of a hypothetical hard sphere that diffuses with the same speed as the particle under examination. However, macromolecules in solution are non-spherical, highly dynamic structures. As such, Hr is indicative of the apparent size of the particles in solution.

DLS was used to probe the formation and quantify the apparent dimension of peptide or peptide/DNA molecular complexes in solution.

2.4. Model membranes

Cellular membranes are highly dynamic and very heterogeneous systems composed of several kinds of lipids and proteins. In fact, the fluid mosaic model had to be continuously readapted to fit the most recent findings in membrane composition and organization (145). For instance, it is now recognized that each organelle has particular lipid content, and that lipid-lipid and lipid-protein associations occur by the formation of specialized regions. Membrane asymmetry is also functionally important. For example, in most eukaryotic cells the exoplasmic leaflet of the plasma membrane is essentially comprised of phosphatidylcholine (PC), the most predominant phospholipid in mammalian membranes, and some sphingolipids, whereas the cytoplasmic leaflet is enriched in phosphatidylethanolamine (PE) and phosphatidylserine (PS) (146, 147).

Due to the great complexity of biological membranes, simple artificial membrane models have been used in the study of biophysical phenomena. Different classes of biomimetic membranes have been developed but lipid vesicles, also known as liposomes, are among the most commonly used. Unillamelar vesicles are well-defined, mono-disperse colloids comprised of a single lipid bilayer folded into a closed spherical structure (146-149). These vesicles can be obtained with different diameters, ranging from 50 nm (small unilamellar vesicles – SUVs), 100 nm (Large unilamellar vesicles – LUV) to giant unilamellar vesicles of 400- >1000 nm diameter.

These models have limitations such as the lack of an extracellular matrix or proteins, and incapacity to emulate energy-dependent processes (e.g. endocytosis). However, they have low cost, are readily prepared, and most importantly can be assembled from a great variety of commercially available lipid combinations, and thus can be tailored to approach any given type of membrane/lipid

domain (147, 148). As such they are extensively employed in biophysical assays of peptide-lipid interactions, which are highly relevant in MAP research. The lipids 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) and and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-L-serine (POPS) were employed in preparations throughout the work described in this document. Due to its abundance in eukaryotic cells, POPC is employed in biophysical assays as a standard for emulating mammalian membranes (118). However POPC is a zwitterionic lipid, and therefore anionic POPS, also present in mammalian cells, was introduced in POPC/POPS (4:1) vesicles to study the effect of charge on peptide activity.

.

3. Materials and methods

3.1. Computational method for identifying CPPs in structural viral proteins

The primary objective of this work was to identify regions displaying CPP potential in structural proteins from viruses. To search for these CPP regions, viral structural proteins had to be identified, their amino acid sequences had to be obtained, and putative CPP sequences had to be located inside those sequences. To achieve these goals an assortment of *in silico* tools were used, including the ViralZone(105) and CellPPD(126) websites and the CLANs software package (127).

Using the Expasy Viralzone website, a large sample of annotated structural proteins from viruses was located. The amino acid sequences of each protein was collected from the UniProtKB database (128). These 270 amino acid sequences from structural proteins of viruses (capsid, envelope and membrane proteins) belonging to every branch of the Baltimore system were uploaded and scanned in the CellPPD server (126). Putative CPP sequences of 20 amino acids in length were screened. All sequences produced by CellPPD for each viral protein were ordered according to their SVM score, and all those predicted as non-CPPs (SVM score < 0) were excluded. Amino acid sequences scoring positive SVM values were collected and renamed as viralCPPs. About 2400 CPP candidates were retrieved from this process.

To select a manageable set of viralCPPs for experimental validation and exclude iterative CPP sequences, the 2400 viralCPP sequences were submitted to CLANS (127), a software package for visualizing protein clusters based on pairwise sequence similarity. CLANS produced 35 clusters of similar viralCPP sequences which were ranked according to their assigned CellPPD SVM score. The top scoring 5% were selected from each cluster, reducing the original set of 2400 viralCPPs to 150 sequences. From this group, a final collection of 14 viralCPPs (Table 3.1 and Annex A) was manually selected for synthesis and experimental testing. The criteria for this manual selection consisted in obtaining sequences with distinct biochemical properties and covering a wide range of SVM scores (0.06-1.63). Each of the 14 viralCPP sequences were blasted against currently annotated CPP sequences using a tool integrated in CPPsite, to assure that these peptides were original and distinct to CPPs currently in that database.

3.1.1. In silico analyses of viralCPP sequences

Computational analyses of viralCPPs (Annex A) were conducted using several online tools. PSIPRED (150) was used to predict viralCPP propensity for formation of secondary structures. Heliquest (151) was employed in the construction of helical projections and to infer about the

amphipathicity of each peptide sequence. Illustrative secondary structure models were computed by PEPstr (152) and Monte Carlo simulations were ran on MCpep server (153) for simulation of viralCPP helical content percentage in aqueous and lipid membrane environments.

3.2. viralCPP synthesis and reconstitution

The 14 viralCPPs were commercially obtained from JPT Technologies (Berlin, Germany), having been synthesized by Fmoc solid-phase chemistry to a purity of >90%. All lyophilized peptides were reconstituted in Milli-Q water (MerkMillipore Integral Water Purification System). Aliquots were prepared at a concentration of 500 μM and stored at -20ºC.

3.3. Chemicals

All ionic salts used in this work, namely NaCl, KCl, KH_2PO_4, Na_2HPO_4, NaF and N-2-Hydroxyethylpiperazine-N′-2-ethanesulfonic Acid (HEPES), as well as ethanol and spectroscopic grade chloroform, were obtained from Merk (Darmstadt, Germany). The lipids POPC and POPS, where purchased from Avanti Polar Lipids (Alabaster, AL, USA). The potentiometric probe di-8-ANEPPS was attained from Sigma-Aldrich (St. Louis, MO, USA). Fluorescent dyes Hoechst 33342 and CellMask Deep Red, GlutaMAX Dulbeco's Modified Eagle Medium, optiMEM (reduced serum media) , Fetal Bovine Serum (FBS), Penicillin-Streptomycin antibiotic solutions, Tryple Express (commercial grade trypsin solution), as well as the stock of ssDNA labeled with Alexa Fluor 488 and the Live/dead viability/cytotoxicity kit for mammalian cells, were purchased from Life Technologies (Carlsbad, CA, USA),

3.4. Delivery of ssDNA into cells by viralCPPs

3.4.1. Cell culture procedures

Human Embryonic Kidney cells (HEK 293) were used to conduct FS and FC analyses, while Baby Hamster Kidney fibroblasts (BHK 21) were employed for confocal microscopy observations. Both cell lines were obtained from ATCC (Manassas, VA, USA). Cells from both lines were cultured using the same procedures. Cryo-preserved cell vials stored at -80ºC in liquid nitrogen were quickly thawed in a heated bath at 37ºC immediately prior to culture initiation. The cells were then transferred to Thermo Scientific (Waltham, MA, USA) Nunc™ EasYFlasks and supplied with fresh GlutaMAX Dulbeco's Modified Eagle Medium enriched in glucose and

pyruvate, and supplemented with 10% v/v FBS and 1% v/v Penicillin-Streptomycin antibiotic solution (10 000 U/ml). A growth period of 2-3 days was allowed for cell expansion before confluence. A maximum of 4 passages was permitted, to ensure cell viability and phenotypic consistency. All cell handling was conducted in a laminar flow hood sterile environment according to good laboratory practice standards. A humidified incubator with 5% CO_2 atmospheric content at 37ºC was used for cell storage and growth. Cell counting and viability assessments were made in a Millipore (Billerica, MA, USA) Scepter 2.0 Automated Cell Counter using 60 µL disposable sensor chips.

3.4.2. Incubation of viralCPP:ssDNA conjugates with HEK cells

Aliquots of each viralCPP were thawed at room temperature and prepared in small eppendorf tubes to be tested at final concentrations of 1, 5 and 10 µM. A solution of ssDNA oligonucleotides (ACGTGCTGAGCCTAC) labeled with Alexa Fluor-488 fluorescent probes (ssDNA-A488) was added to each viralCPP aliquot to be assayed at 0.1 µM. Each sample was mixed by short spin in an eppendorf (Hamburg, Germany) MiniSpin centrifuge prior to cell addition, to promote formation of electrostatically bound viralCPP and ssDNA-A488 conjugates (viralCPP:ssDNA-A488).

HEK 293 cells were washed in standard Phosphate Buffer Saline (PBS), pH 7.4, and harvested by trypsinization (2 minute treatment with Tryple Express). Fresh culture medium was added to inactivate trypsin activity and to collect the cells. Cell pellets were produced by centrifugation at 1250 RPMs for 5 minutes in an eppendorf 5810-R centrifuge, washed twice in PBS and resuspended with optiMEM.

Collected cells were added to the previously prepared viralCPP:ssDNA-A488 aliquots at a concentration of 7.5×10^5 cells/mL to a final volume of 500 µL. Incubation proceeded for 1 hour in a VWR Incubating Orbital Shaker, at 37ºC and under gentle mixing (100 RPMs). Following incubation, pellets were obtained from each sample by centrifugation at 2000 RPM, for 5 minutes at room temperature in a Thermo Scientific Heraeus Multifuge 1L-R. The supernatant was discarded to remove signal from non-internalized peptide:ssDNA-A488, followed by re-suspension of pellets in 500 µL of PBS. Controls for the experiment were equally prepared. These included a sample of cells and cells incubated with ssDNA-A488 but no peptide. The signal of A488 probe retained per sample was then analyzed by FS, FC and by a microplate reader (PR).

3.4.3. Fluorescence spectroscopy

An Edinburgh Instruments (Livingston, UK) FS920 Fluorescence Spectrometer, equipped with 2 double monochromators and a 750 W xenon lamp, was employed for obtaining high-sensitivity fluorescence intensity spectra from each sample prepared as previously described in section 3.4.2. Briefly; 100 µL from each of these solutions were diluted 4:1 with PBS. Hellma Analytics (Müllheim, Germany) 5x5 mm quartz cuvettes were used to collect the emission spectrum of A488 (Figure 2.2 and Annex B1) retained in the samples. Emission scan readings were performed at wavelengths ranging from 500 to 650 nm in 1 nm steps with a dwell time of 0.2 seconds. Three acquisitions were conducted per sample. The excitation wavelength was fixed at 488 nm and both excitation and emission slits were set at 10 nm. Background was subtracted from every emission spectrum. The integral of every spectrum was calculated to quantify total fluorescence intensity. The results were plotted as bars displaying the mean value of five replicates and the standard error of the mean (SEM).

3.4.4. Microplate fluorescence reading and flow cytometry

In addition to FS, FC and PR measurements were conducted with the samples descried on section 3.4.2. Two volumes (technical replicates), of 200 µL each, were distributed per sample into the wells of a fluorescence ready V-bottom Eppendorf Polypropylene 96-Well Microplate.

Fluorescence measurements were recorded by a Tecan (Maenedorf, Switzerland) Infinite F500 multimode Microplate Reader in fluorescence top reading mode. Excitation wavelength was defined with a 485 nm filter (20 nm bandwidth), while emission wavelength was set with a 535 nm filter with 25 nm bandwidth.

Immediately after PR data collection, the same 96-well plate was measured by FC analyses. Fixed volumes of 180 µL per sample were injected into a BD (San Jose, CA, USA) LSRFortessa cell analyzer equipped with a high throughput screening module and a blue 488 nm laser. The FL 530/30 BP channel was used to acquire A488 fluorescence intensity signal as well as FSC and SSC signals for each event recorded.

Background subtraction was performed before plotting of PR and FC data. PR and FC A488 fluorescence intensity data were plotted as bars displaying the mean value of five replicates and SEM (Annex B.2.3). To quantify the percentage of cells loaded with CPP:ssDNA-A488 in FC measurements, positive events were gated using FlowJo v8.7 software, and the ratio of positive

events to total number of counts was calculated and plotted as bars displaying the mean value of five replicates and SEM.

3.4.5. Effect of trypsin on viralCPP:ssDNA-A488 delivery

To evaluate the amount of signal obtained from viralCPP:ssDNA-A488 adsorbed at the outer leaflet of cell membranes (i.e. retained in samples but not internalized by cells), each sample prepared and incubated as previously described in section 3.4.2 was treated with Tryple Express. However, cells were resuspended in PBS instead of optiMEM to avoid inhibition of protease activity by the culture media. After the initial incubation with viralCPPs at 5 μM and ssDNA-A488 at 0.1 μM, 5 μL of Tryple Express were added to each sample. Tryple is a commercial grade trypsin solution, and its formulation has not been disclosed by the manufacturer, therefore 5 μL were used as the maximum volume that did not compromise cell viability. Incubation with Tryple proceeded for 1 minute, at 37ºC, in the same orbital shaker incubator as before. After washing, the samples were analyzed by FS, as described in section 3.4.3.

3.4.6. Confocal laser scanning microscopy of CPP:ssDNA-A488 cell internalization

For confocal microscopy (CM) observations of viralCPP:ssDNA-A488 *in vitro* delivery, BHK 21 cells were seeded at 2×10^5 cells/mL in iBidi μ-Slides VI plates and grown overnight. Staining solutions were prepared, namely Hoechst 33342 at 5×10^{-2} mg/mL and Cell Mask Deep Red at 1.5×10^{-3} mg/mL. Each well was then gently washed in PBS, and the cells were stained with 100 μL of Hoechst or with a combination of Hoechst and CellMask Deep Red. After 30 minutes of incubation, the cells were again washed in buffer and 100 μL of a solution of peptide (50 μM) and ssDNA-A488 (2μM) were added to each well, followed by 1 hour of incubation. After replacing the supernatant with fresh PBS, the samples were observed through an inverted confocal point-scanning Zeiss (Jena, Germany) LSM 510 META microscope equipped with Diode 405-30, Argon2, DPSS 561-10, HeNe 594 and HeNe 633 lasers and a temperature control incubator (37°C) with CO2 supply. Images were collected with a Plan-Aprochromat 63x objective from Zeiss. The collected images were analyzed using ImageJ 1.47v software (154) to remove background noise and create composites from the ssDNA-A488 (green), nuclei (Hoechst-blue) and cell membrane (CellMask Deep Red-red) channels.

3.4.7. Live/Dead assay

Live/Dead viability/cytotoxicity kit for mammalian cells (134) is a two colour assay which can be used with FC to determine the viability of cell populations. The kit contains two fluorophores to stain cells, calcein-AM and ethidium homodimer-1. Calcein fluorescence is activated by intracellular esterase activity which is ubiquitous in live mammalian cells. Therefore, calcein is employed to discriminate live cells through green fluorescence staining. On the other hand, ethidium homodimer-1 is cell-impermeable and labels dead cells with red fluorescence by binding to nuclei acids upon loss of plasma membrane integrity. The Live/Dead viability/cytotoxicity kit for mammalian cells was used to evaluate the cytotoxicity of viralCPPs on HEK 293 cells. Supplier's indications (134) were followed to prepare staining samples of calcein-AM and ethidium homodimer-1 for FC analysis. Sample solutions were prepared as described in section 3.4.2 at viralCPP concentrations of 5µM and 10 µM. Controls consisted of live cells, and of dead cells exposed to 70% isopropanol. Fluorescence measurements were conducted in the same cytometer as in previous cases, using a wavelength of 488 nm for excitation and measuring green fluorescence emission for calcein (530/30 bandpass) and red fluorescence emission for ethidium homodimer-1 (610/20 bandpass). Live and dead populations were plotted and gated as references using FlowJo v8.7 (Annex B.2.3) so that percentage of viability could be calculated.

3.5. Biophysical assays

3.5.1. Preparation of lipid vesicles

Lipid vesicles were prepared according to standard procedures described elsewhere (148). Briefly, the desired amounts of each lipid were individually weighted, dissolved in chloroform and mixed in a round bottom flask. The organic solvent was then evaporated via a flux of nitrogen gas and stored overnight under vacuum. The lipid film was rehydrated in buffer (HEPES 10mM, NaF 40mM, pH 7.4) and subjected to 8-10 heating/thawing cycles to promote formation of large multilamellar vesicles. LUVs with the desired diameter were obtained using an Avestin LiposoFast-Basic extruder with a 100 nm sized pore Whatman (Maidstone, UK) Nuclepore membrane filter. For SUV preparation the rehydrated films were submitted to power sonication in a Sonics&Materials VibraCell for 5 minutes.

3.5.2. Di-8-ANEPPS membrane dipole potential sensing assay

For membrane dipole potential sensing procedures the fluorescent probe di-8-ANEPPS was directly added to lipid/chloroform solutions during LUV preparation. In this case, solutions of 200µM of POPC and POPC/POPS (4:1) were obtained stained with 0.5% (n/n) of di-8-ANEPPS. ViralCPPs were initially tested at 10 µM in each lipid system. To assay peptide induced alterations in membrane dipole potential, excitation spectra of di-8-ANEPPS were collected before and after viralCPP addition to the model membranes, using an Edinburgh Instruments FS920 Fluorescence Spectrometer. These spectra (Annex B.4.1) were recorded between 350 nm and 550 nm, in 1 nm step single runs of 3s dwell time. Emission wavelength was set at 606 nm with a 3 nm and 8 nm slits for excitation and emission respectively. Each spectrum was corrected for background and normalized by its underlying area. Sinusoidal differential spectra were obtained (Figure 2.4 and Annex B.4.3) by subtracting the normalized spectra collected before and after viralCPP addition. In viralCPPs displaying an affinity for model membranes, incremental concentrations were tested to perform dual wavelength ratiometric measurements (137). In these cases, excitation spectra were collected at viralCPP concentrations of 1, 5 and 10 µM (Annex B.4.2) and differential spectra were produced as before (Annex B.4.4). **R** was calculated for each of these cases as the ratio of fluorescence intensities (from excitation spectra) at wavelengths where peaks are observed in the differential spectra. Ratio values were normalized for the initial value in undisturbed di-8-ANEPPS LUV spectrum ($\mathbf{R}_0$). $\mathbf{R}/\mathbf{R}_0$ was then plotted as mean values of three replicates.

3.5.3. Circular dichroism spectroscopy

The assessment of viralCPP secondary structure was attained through CD using POPC/POPS (4:1) vesicles. CD spectra were acquired with a JASCO (Easton, MD, USA) spectropolarimeter model J-815, at 25°C in the 200–260 nm range, using Hellma Analytics 0.1 cm quartz cells. ViralCPPs were first assayed at 100 µM in HEPES buffer, in the absence or presence of LUVS at incremental concentrations of 0, 0.5, 1, 2, 4 and 6 mM. Each final spectrum corresponded to the average of 10 runs, which were subsequently corrected for buffer or LUV baselines. Peptides were also assayed under the same conditions using SUVs at concentrations of 0, 0.5 and 1 mM. All spectra were plotted as a function of mean residue ellipticity, and submitted to the K2D3 web server (155) for assessment of content in standard secondary structure elements.

3.5.4. Dynamic light scattering

DLS experiments were carried out to assay the hydrodynamic diameter of viralCPPs and viralCPP:ssDNA-A488 complexes. These measurements were conducted on a Malvern (Malvern, UK) Zetasizer Nano ZS with backscattering detection at 173°, equipped with a He-Ne laser (632.8 nm) at 25°C , following 10 min of equilibration, using Hellma Quartz cuvettes with a 3 mm light path and center 9.65 mm. Each replicate consisted of 15 measurements, and 2 replicates per sample were conducted and averaged. ViralCPP (100μM) and viralCPP(100 μM):ssDNA(20 μM) solutions were then prepared in HEPEs and assayed. To assure complete removal of any foreign particle contaminants, every solvent was passed through a 0.45 μm syringe filters from VWR International. DLS size/number distribution tables were extracted from the Malvern Zetasizer software v7.02. Mean hydrodynamic diameter values were averaged from the results of 2 replicates and plotted as histograms (Annex B.5.1 and B.5.2).

4. Results and discussion

The foundation of this project was provided by several factors. The first was the rational that structural viral proteins exhibit a natural affinity for biological membranes. Another issue was that CPPs from viral origin have already been described. However, despite these two facts, viruses remain an underexplored source of MAPs. Also critical was the identification of two regions in DENV C protein which were isolated as potent CPPs, namely pepR and pepM (118, 156). These findings led to the hypothesis that structural proteins from viruses are a rich and currently underexplored source of CPPs. Under this hypothesis, the work herein described had several objectives. The first objective was the identification of novel CPP sequences, termed viralCPPs, from putative regions in structural proteins belonging to a wide and diverse sample of viruses. A following goal was that of experimentally validating a set of viralCPPs *in vitro* as potential vectors for gene therapy. The last objective was to conduct preliminary biophysical studies to characterize this set of viralCPPs in order to understand their efficiency in the delivery of molecular cargo.

To fulfil these objectives the work was divided into three stages. The first stage concerned the prediction of viralCPP sequences with the aid of computational tools. Therefore structural viral proteins were identified, and their amino acid sequences were collected, through the ViralZone directory. These were, in turn, submitted to CellPPD, an online server for CPP prediction and for the identification of probable CPP regions in proteins. 2400 viralCPP sequences were obtained from this procedure, each a possible CPP lead. A manageable set of 14 viralCPPs was selected from this pool, and synthesized, to be experimentally tested. Thus, in the second stage of the work viralCPPs were conjugated with ssDNA-A488 oligonucleotides and were assayed as DDSs for gene therapy. HEK 293 cells were incubated with viralCPP:ssDNA-488 and uptake efficiency was evaluated by the fluorescence signal of the A-488 probe. This signal was measured by FS and FC. During this stage other assays were also conducted. Trypsin was used to assess the adsorption of probe to cell membranes; confocal microscopy allowed the visual confirmation of viralCPP-mediated delivery of cargo to BHK 21 cells; and a toxicity study was conducted with a Live/Dead cytotoxicity FC kit. In the final stage of this work, preliminary biophysical methodologies were employed to assess physical-chemical properties of viralCPPs and viralCPP:ssDNA-A488. In this later section membrane dipole potential sensing by di-8-ANEPPS was used to study the propensity of each viralCPP towards model membranes. CD was employed to determine the formation of secondary structures by viralCPPs in the presence of lipid membranes. Finally the formation of CPP/ssDNA conjugates was relayed by shifts in hydrodynamic diameters evaluated by DLS spectroscopy.

The results obtained from each of these procedures as well as their interpretation, are described in this chapter.

4.1. ViralCPPs: identifying novel CPP sequences in structural viral proteins

The current section presents results from the first stage of the work, namely the identification of viralCPP sequences. As previously mentioned, viralCPPs are novel amino acid sequences with CPP potential identified in structural viral proteins through computational methodologies. Upon a process of selection from 2400 CPP candidates, a final set of 14 viralCPP sequences was obtained for experimental work. These sequences are displayed in Table 4.1, along with their SVM score, positive net charge value, virus of origin and protein source.

This set of viralCPPs was chosen based on several criteria. The first criterion was promoting biochemical diversity between amino acid sequences. This is evident in the column of Table 4.1 displaying the amino acid sequences of each peptide. In general, the 6 viralCPPs on the top of the column (viralCPPs 1415-0979) have lower content in basic amino acid residues and are more varied in terms of constitution, whereas the remaining sequences (viralCPPs 0417-0769) are more uniform and mainly rich in arginines and lysines. However, no two sequences are identical.

A second criterion was that of covering a variety of SVM scores. These ranged from a modest 0.06 attributed to viralCPP 0956 and viralCPP 0554, to the highest value of 1.63 associated with viralCPP 2319. By standard definition of the CellPPD algorithm, positive SVM scores signify CPP potential, and higher-scoring SVM hits are more probable CPPs. As the great majority of known CPP sequences are rich in basic amino acid residues, this trend was necessarily reproduced in the learning sets that instructed the machine learning algorithm of CellPPD (126). Thus, high SVM scores are related to abundance in arginine and lysine content. In fact, presenting a poly-arginine sequence to CellPPD produces the highest SVM score of 2.6, while the Tat peptide is predicted from the amino acid sequence of the TAT protein with an SVM score of 1.3. This bias of CellPPD towards basic residues is equally observable in the set of viralCPP sequences. The 8 viralCPPs on the bottom of Table 4.1 (viralCPPs 0417-0769) are particularly rich in R and K residues, exhibit higher positive net charge magnitude, and collectively gather the highest SVM values. Conversely, the remaining viralCPPs are more unusual relative to the landscape of currently known CPPs by having a lower percentage of basic residues, and therefore present much lower SVM scores. This however, justifies their relevancy as study subjects, and also allows that they be used as controls for assessing the sensitivity of CellPPD.

Table 4.1 – viralCPPs: sequences, SVM score, charge, viral and protein origin

viralCPP ID	Sequence	SVM score	Positive Net Charge	Virus of Origin	Protein Source
1415	DLRTNPPPTEPPSRKSKLMS	0.07	2	Indian citrus ringspot virus	Capsid
0956	GHWSRDCTQPRPPPGPCPLC	0.06	1	Human T-cell leukemia virus 1	Capsid
1396	PIRVKRPKKPIAKRNSNISR	0.09	8	Influenza C virus	Capsid
1203	PTTKTTNKRDPKTPAKTTKK	0.08	6	Human respiratory syncytial virus B	Envelope
0554	RNLFRVPKYINGTKLKNTMR	0.06	6	Cytomegalovirus	Envelope
0979	LPPFSLAPVPPPATRRRRAV	0.06	4	Human T-cell leukemia virus 2	Envelope
0417	SPRRRTPSPRRRRSQSPRRR	0.93	10	Hepatitis B virus genotype C	Capsid
1779	GRRGPRRANQNGTRRRRRRT	0.83	10	Barley yellow dwarf virus	Capsid
0275	KKRYKKKYKAYKPYKKKKKF	0.73	13	Cauliflower mosaic virus	Capsid
2319	WRRRYRRWRRRRRWRRRPRR	1.63	15	Torque teno douroucouli virus	Capsid
1721	GKTQQQKKKDKQADKKKKKP	0.63	8	Sindbis Virus	Capsid
0667	RPRRRATTRRRITTGTRRRR	0.94	11	Human adenovirus C serotype 1	Capsid
1754	KQKQQAPQNNTNQKKQPPKK	0.26	6	Chikungunya Virus	Capsid
0769	RRLTLRQLLGLGSRRRRRSR	0.59	9	Fowl adenovirus A serotype 1	Capsid

Another criterion in the selection of the sequences present in Table 4.1 was that of including examples from different CPP families (e.g. amphipathic, cationic, proline-rich). To help categorize the viralCPPs, *in silico* analyses were performed to each amino acid sequence using the tools described in section 3.1.1. The results from these procedures can be viewed in Figure 4.1 and Annex A.

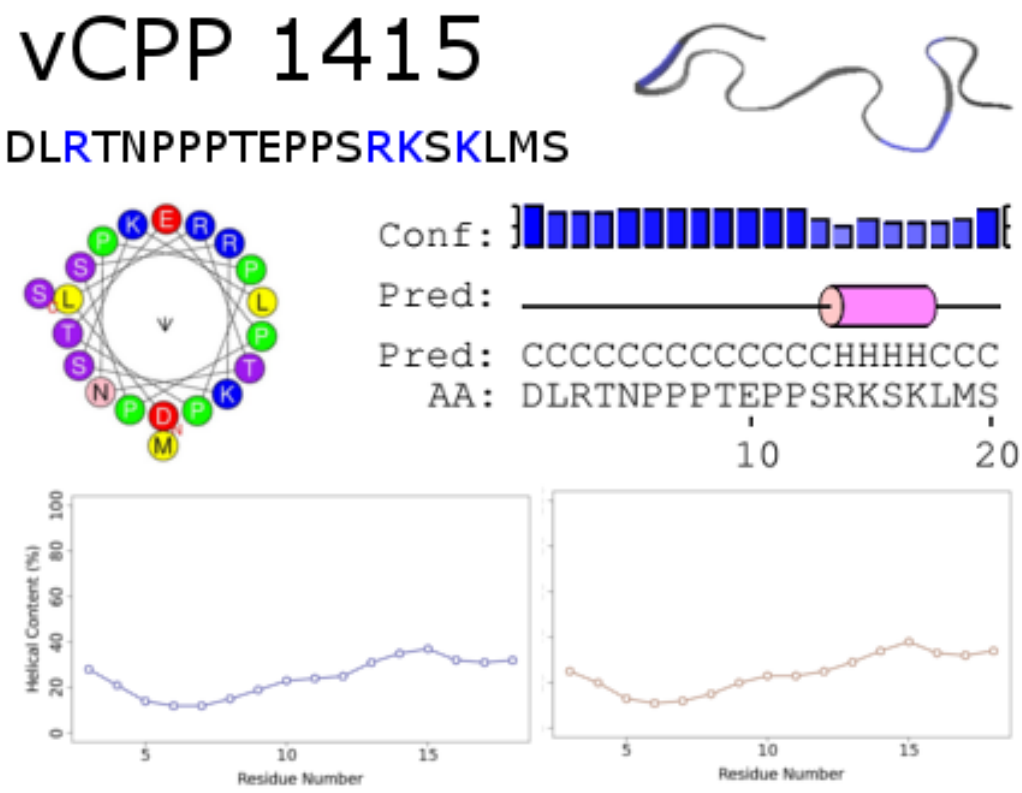


Figure 4.1 – Example of viralCPP *in silico* analysis – The amino acid sequence of viralCPP 1415 is highlighted in blue for identification of basic residues. On the left, a helical projection is illustrated along with the mean amphipatic moment vector. To the right, a tridimensional simulated structure is depicted above a prediction of secondary structure. Helical content predicted using Monte Carlo simulations of the peptide in aqueous (blue curve) and 20% anionic membrane environment (brown curve) are shown below.

Among these results, the helical projections of viralCPPs (Annex A) produced with the Heliquest(151) web tool were used to classify each peptide according to the criteria proposed by Milletti (7) and described in section 1.2.3. ViralCPPs 1415, 1203, 0979 and 1754 can be held as either cationic or amphipathic proline-rich, because they fulfil the minimum density of charge required to be inscribed into the first category (+ 0.2 per residue), but also display considerable content in proline residues. Similarly, viralCPP 0956 can be classified as a proline-rich peptide, although it is also possible that it be viewed as hydrophobic, due to a very low amount of positive charges and the high content in hydrophobic amino acid residues. On the other hand, viralCPPs 1396 and 0554 can be grouped in the secondary amphipathic category due to the opposing distribution of basic and hydrophobic residues in their helical projections. All remaining peptides can be ascribed to the cationic family, due to very high content in basic amino acids. However, in this subset, viralCPP 0417 can also be considered proline-rich, and there is evidence to support

classification of viralCPPs 0275, 2319 and 0769 as secondary amphipatic α-helical peptides. Although ambiguous, this categorization effort confirms that representatives of different CPP families were indeed present in the viralCPP set. Removing these ambiguities would imply that further data concerning viralCPP structure and activity be collected. This biophysical data would be used to infer how the biochemical properties of each viralCPP are related to its activity. For instance, the structure of every peptide could be determined in aqueous and membrane environments through techniques such as CD or nuclear magnetic resonance (42). This would be important because secondary conformations are related to amphipathicity, and therefore to CPP activity. Furthermore, studying the effect of amino acid substitutions on residues which are suspected to play a key role in viralCPP membrane interaction or uptake could also clarify this issue (157).

Another attribute worth stating in the set of viralCPPs is the variety in viral origins. Some of these peptides were identified in intensely studied animal (including human) viruses, while the set also included viruses from poorly characterized *taxa*. Three plant viruses also served as origin for viralCPPs 1415, 1779 and 0275, namely the indian citrus ringspot virus, the barley yellow dwarf virus and the cauliflower mosaic virus. No unifying trend could be distinguished in the overall morphology of these viruses. The set contained enveloped and non-enveloped virions, some of simple morphology (simple capsid surrounding an oligonucleotide), of distinct RNA and DNA genomes and uptake mechanisms. Finally, it can be stated that the majority of these viralCPPs were withdrawn from sequences of Capsid proteins, while Envelope proteins were a minority, and no examples from Membrane proteins were included.

4.2. Validation of viralCPPs as DDSs for gene therapy

4.2.1. Delivery of viralCPP:ssDNA-A488 into HEK cells evaluated through FS and FC

The second stage of this work had the objective of validating the set of viralCPPs *in vitro* to confirm their activity as CPPs. A protocol was devised for this end, consisting on assaying viralCPP-mediated delivery of ssDNA oligonucleotides labeled with a fluorescent probe (A-488) though FS and FC. The first step of this protocol was to conjugate each viralCPP (at 1, 5, and 10 μM) with the oligonucleotide probes (at 0.1μM) via formation of electrostatically bound complexes. These conditions allowed that a range of viralCPP:ssDNA-A488 molar ratios - 10:1, 50:1 and 100:1 – be tested. Each sample of viralCPP:ssDNA-A488 was then incubated with HEK 293 cells. FS was used to collect A-488 emission spectra and relay the efficiency of probe delivery to each sample

through the quantification of probe fluorescence intensities. On the other hand, FC was used to quantify the percentage of positive delivery events in the total population of cells per sample. Supplementary fluorescence measurements were also conducted with FC and a PR and are provided in Annex B2.

PepM was chosen as a positive control for these experiments because it is a CPP derived from a structural viral protein (C protein of DENV). Moreover, its intracellular delivery capabilities were recently characterized, and the peptide demonstrated excellent cellular uptake into several cell lines when conjugated with ssDNA cargo identical to the one used in this protocol (117, 118).

Figure 4.2 gathers data from FS (upper panel) and FC (lower panel) with respect to viralCPP-mediated delivery of ssDNA oligonucleotides *in vitro*.

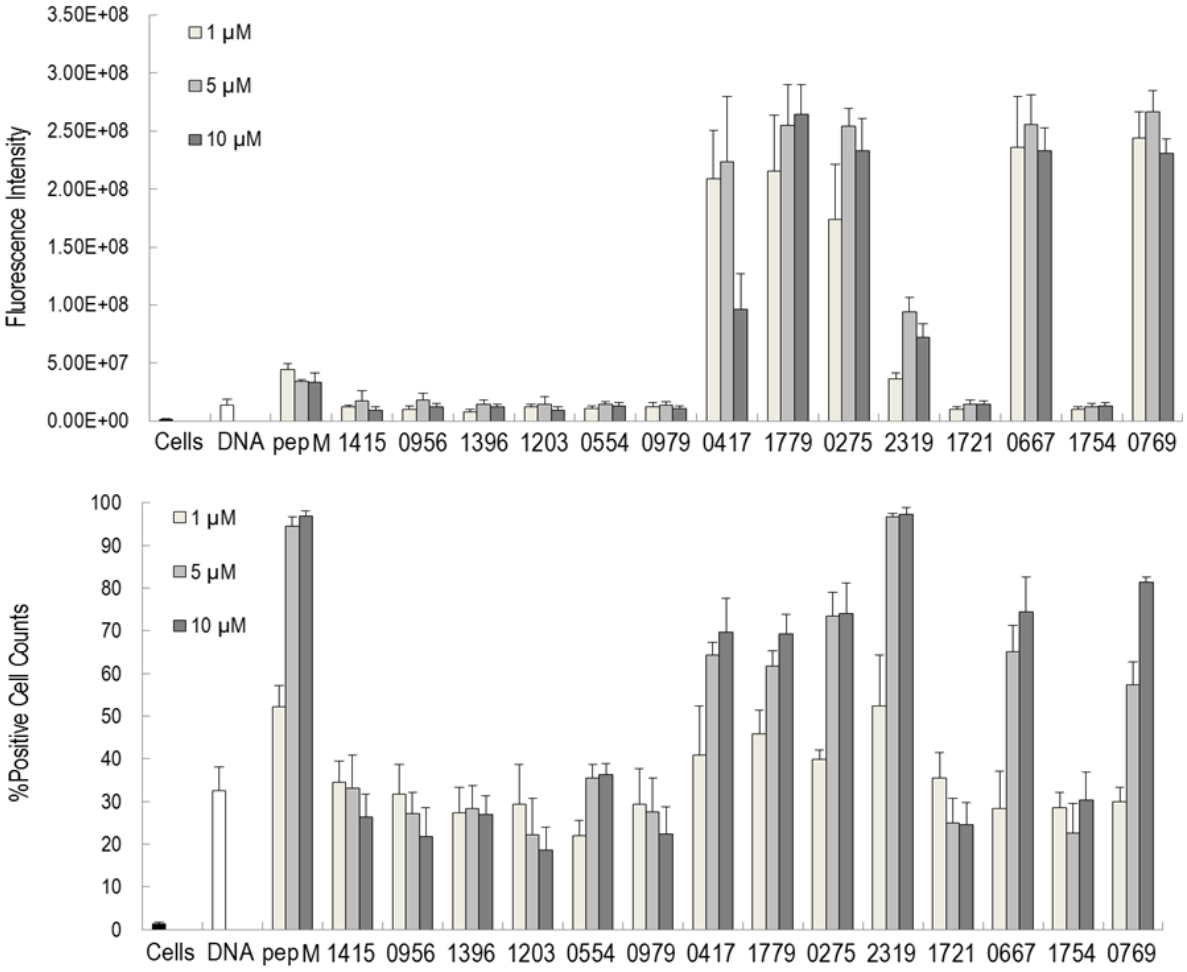


Figure 4.2 – Determination of viralCPP efficiency by FS and FC – Peptide concentrations are depicted in grey scale, ssDNA controls in white and cell controls in black. (Upper panel) FS Intensity of A488, n = 5 with SEM error bars; (Lower Panel) Percentage of positive cell counts by FC, n = 5 with SEM error bars.

Both panels of the figure show that cell auto-fluorescence was negligible. The upper panel of Figure 4.2 demonstrates that controls of cells incubated with ssDNA-A488 produced low signal intensity, signifying that the probe could not be efficiently retained by cells. In contrast, a mean of up to 40% in positive cell counts was observed in FC measurements of ssDNA controls, as displayed on the lower panel of Figure 4.2. This value indicates that a residual amount of the oligonucleotide probe was retained in the samples even in the absence of viralCPPs. Nonetheless, this residual signal did not obscure the determination of a viralCPP efficiency profile. Indeed, both panels of Figure 4.2 show that viralCPPs 1415, 0956, 1396, 1203, 0554, 0979, as well as 1721 and 1754, did not improve ssDNA-A488 delivery into HEK 293 cells. As such, in this particular set of conditions, there was no evidence to validate them as CPPs. Conversely, viralCPPs 1779, 0417, 0275, 2319, 0667 and 0769 greatly augmented the signal of ssDNA probe retained in each sample, indicating CPP potential in these peptides. Based on these results the set of 14 viralCPPs could be divided into 2 groups. The first group was comprised by the 8 peptides which are not CPPs, whereas the second group consisted of the 6 viralCPPs validated as CPPs.

Data form the upper and lower panels of Figure 4.2 also showed that increasing peptide concentrations had a positive effect on delivery efficiency. This was true for every CPP with the exception of viralCPP 0417. This peptide experienced diminished FS intensity at 10 μM, while retaining higher ratio of FC positive cell counts relative to 1 μM and 5 μM. The opposite was true for viralCPP 0769 at 5 μM, the concentration at which fluorescence intensity was most potentiated, but at the cost of a 20% reduction in positive cell counts. Otherwise, every viralCPP produced nearly identical efficiency at 5 μM and 10 μM. Therefore, using viralCPPs at 5μM would be a better compromise in terms of potency/dosage relationship.

It can be seen on the upper panel of Figure 4.2 that pep M's FS intensity was 2-3 times that of the negative sample. However, this effect was much smaller than that of viralCPPs surpassing 10-fold the control's fluorescence intensity. Moreover, pepM has been described as highly efficient in oligonucleotide delivery (118), a finding which was supported by the FC evidence present in Figure 4.2's lower panel. The success rate of pepM ssDNA delivery was nearby 100%, whereas for viralCPPs 0417, 1779, 0275, 0667 and 0769 positive cell count was about 20% lower. On the other hand, viralCPP 2319 constituted a very similar case to pepM. This peptide exhibited considerable increase in fluorescence intensity relative to control samples, yet a much lower increase than that of the other five CPPs. Additionally viralCPP 2319 also demonstrated nearly 100% positive cell counts for ssDNA-A488 internalization. Taking into account the evidence of both panels in Figure 4.2, as well as supplementary PR and FC data in Annex B2, it can be reasoned that pepM and

viralCPP 2319 were in fact the most efficient CPPs for delivery of ssDNA-A488 into HEK cells. Therefore, pepM and viralCPP 2319 efficiency in ssDNA-A488 delivery might have been underestimated by FS measurements, possibly due to phenomena resulting in quenching of A488 fluorescence.

Correlation between viralCPP SVM scores and delivery efficiency

Figure 4.3 shows the correlation between the SVM score of each viralCPP and the results from FS and FC measurements. This data was used to calculate Pearson correlation coefficients, displayed in Table 4.2.

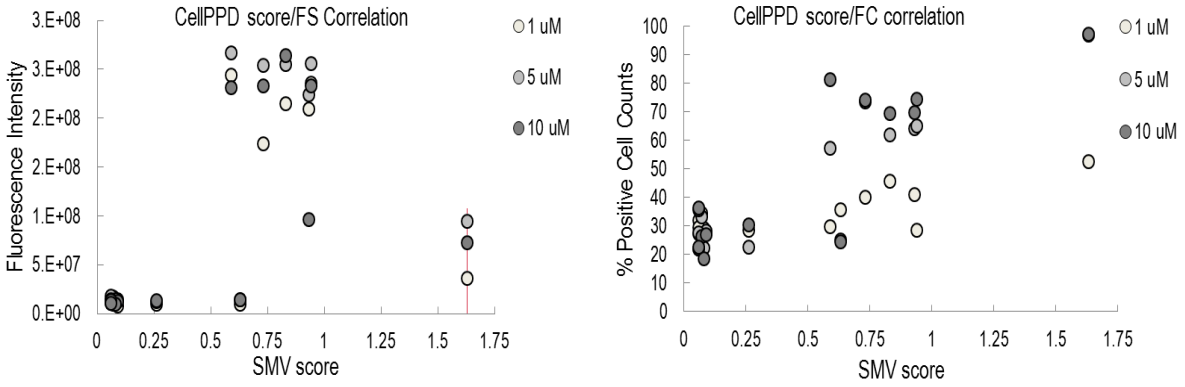


Figure 4.3 - Correlation between viralCPP SVM scores and efficiency – (Left panel) Scatter plot illustrating correlation between CellPPD SVM scores and FS results; (Right panel) Scatter plot illustrating the correlation between CellPPD SVM score and FC results.

Table 4.2 - Pearson correlation coefficient of viralCPP SVM scores and efficiency

	Pearson Correlation Coefficient		
	Fluorescence Intensity		% Positive Cell Counts
	Uncorrected	Corrected*	
1µM	0.54	0.87	0.79
5µM	0.62	0.86	0.90
10µM	0.55	0.79	0.88

*exclusion of viralCPP 2319 data points for more accurate estimation of Pearson correlation coefficient

For FS data in the left panel of Figure 4.3 and in Table 4.2, Pearson correlation coefficients (~ 0.5-0.6) provided minor support of positive correlation. Figure 4.3 indicates that this was due to the contribution of the three data points belonging to viralCPP 2319, corresponding to the highest SVM scores (red line crossing coordinate 1.63 of the horizontal axis). It was referred in section 4.2.1 that the efficiency of viralCPP 2319 as a DDS might have been underestimated by FS measurements. Therefore a correction was introduced to this data by rejecting viralCPP 2319 from the calculation of Pearson's coefficient. This correction resulted in improved estimates of correlation (> 0.75) between viralCPP SVM score and efficiency. As for FC data, very good indicators of correlation were obtained (~0.9). This analysis indicates that CellPPD scores may have predictive value in terms of CPP efficiency.

Indeed, viralCPPs with the lowest SVM scores (< 0.1) were eliminated as CPP candidates, as previously discussed (Table 4.1and Figure 4.2). Presumably, low positive net charge was the underlying factor for the exclusion of these peptides as CPPs. Apart from low net charge, other determinants to explain poor viralCPP delivery might include insufficient/excessive hydrophobicity (158) or lack of structural flexibility (42). The exception among the group of low scoring viralCPPs was viralCPP 1396. This peptide provides a good example of the difficulty in understanding the biochemical requisites for a CPP. It has substantial net charge (+ 8) and its helical projection (Annex A) suggests that it can acquire considerable degree of amphipathicity, two factors which are shared by a great number of CPPs. However, despite these seemingly favourable traits, viralCPP 1396 displayed low SVM score and poor delivery efficiency. Understanding this fact would require extensive assaying of viralCPP 1396, as the biochemical requisites for CPP activity are still not fully characterized. In fact, shuffling CPP sequences or substituting a single amino-acid residue can often reduce or even abolish CPP activity completely (157). This indicates that single residue positions can be as important to CPP activity as net amino acid content. Two other peptides, namely viralCPP 1721 and 1754 were also excluded as CPPs, although exhibiting considerably higher SVM scores, respectively 0.63 and 0.23. Nonetheless their charge magnitude was identical to that of lower scoring viralCPPs, which might once again explain their lack of efficacy.

All viralCPP sequences with acidic residues were found to have reduced CPP potential (viralCPP 1415, 0956, 1203 and 1721), which suggests that these amino acids were disruptive to uptake. For instance, viralCPP 1721 has considerable number of positive charges (10 K residues), but a lower net charge of + 8 due to two negatively-charged aspartic acids. Therefore, the exclusion of viralCPP 1721 as a CPP supports the hypothesis that net charge is a broader and more critical CPP parameter than abundance in positive residues. Electrostatic repulsion between these viralCPPs

and ssDNA-A488 or cell membranes may have occurred due to the acidic residues, hampering viralCPP delivery efficiency.

On the other hand, viralCPPs displaying high SVM scores were rich in R and K residues. These peptides exhibited a density of positive net charge per amino acid ranging between + 0.45 and + 0.75. Furthermore, some of these amino acid sequences (Table 4.1) are comprised from interesting repetitive patterns. For instance, viralCPP 0417 is constituted by repeats of small SPRRR motifs. ViralCPP 0275 is made up by stretches of K residues intercalated with hydrophobic tyrosine. An analogous case happens with viralCPP 2319, displaying basic Rs punctuated by hydrophobic tryptophan residues. The sequence of viralCPP 0769 is also interesting, with hydrophobic leucines on its central part and stretches of R residues at the termini. Current CPP literature postulates that positive residues might be responsible for primary electrostatic interactions with membranes or extracellular matrix components. Additionally hydrophobic residues are thought to promote subsequent insertion of CPPs into biological membranes (52). Therefore these basic and hydrophobic residues are related to CPP activity. This finding is in agreement with the amino acid content in the group of viralCPPs validated as DDSs. Moreover, viralCPP 2319 was more efficient in delivery relative to viralCPP 0275. Both these peptides are similar, but viralCPP 2319 is rich in R whereas viralCPP 0275 has K residues. This in agreement with reports in the literature stating that, of the two basic residues, R increases CPP efficiency relative to K (39). However, viralCPP 2319 and 0275 are not perfect R/K analogues, and therefore differences in activity could also be explained by content and relative position of hydrophobic residues, or to slight differences in net charge magnitude (157).

Advantages and drawbacks of the protocol

The protocol that was developed for testing the delivery of viralCPP:ssDNA-A488 into HEK cells has several advantages. It is a fast, practical, and resource efficient way of assessing CPP potential from a large set of leads. As it was the objective of this work to assess viralCPPs as potential DDSs, ssDNA oligonucleotide cargo was labeled with an A-488 fluorescent probe. By labeling cargo molecules, in place of viralCPPs, it was possible to directly probe the efficiency of cargo delivery, instead of simply assessing peptide internalization. In this particular case, the use of ssDNA-A488 cargo was envisioned as a model to study the transport of nucleotides for gene therapy. Furthermore, probing cargo molecules instead of the peptides avoids the introduction of extrinsic hydrophobic/steric effects in the amino acid chains. As such, biophysical studies can be

conducted on viralCPPs without artifacts due to labeling (159). Moreover, delivery of any other biomolecule could, in principle, be tested just as easily with this protocol, provided that it was adequately labeled and could electrostatically bind to the CPPs. Finally, by assaying viralCPPs at 1μM, 5μM, and 10μM, while maintaining ssDNA-A488 at 0.1 μM, different peptide concentrations and peptide/cargo ratios were simultaneously tested. This was important because both parameters must be tested and optimized to potentiate CPP activity.

However the protocol also has some drawbacks. It can only provide an indirect assessment of CPP potential through measurements of A-488 fluorescence, and thus cannot distinguish membrane adsorption of viralCPP:ssDNA-488 from intracellular uptake. Furthermore, successful therapeutic delivery often necessitates endosomal escape of the DDSs which cannot be resolved through this methodology. However, the protocol could be easily adapted to overcome these limitations. For instance, viralCPPs could be used to transfect cells with a reporter plasmid or with an antisense oligonucleotide for gene knockdown. These strategies could ascertain if the CPPs can promote endosomal escape and functional delivery by exerting phenotypic changes in the cell lines. Furthermore, to remove the signal from cargo absorbed at the outer leaflet of cytoplasmic membranes, a simple modification was introduced to the protocol, as described in section 4.2.2.

Given these considerations, the current protocol should be regarded as a reliable method for preliminary assessment of CPP potential, *i.e.* to rapidly exclude poor CPP leads. After this assessment, further characterization is required to fully confirm viralCPPs as established DDSs. In this regard, CM observations of viralCPP:ssDNA-A488 delivery into live cells were performed, as described on section 4.2.3.

4.2.2. Effect of trypsin on viralCPP mediated ssDNA-A488 delivery

The protocol described in section 4.2.1 has the limitation of accounting for signal from probe retained in the cytoplasmic membrane of cells. To resolve this issue, an additional step was added to its execution. This modification consisted in treating the cells incubated with viralCPP:ssDNA-A488 with trypsin. Trypsin is a serine protease which cleaves peptides at the carboxyl terminus of K and R residues. As such, it was used to hydrolyze viralCPPs which might have been retained in solution or stuck to the outer leaflet of plasma membranes. This treatment removes false positive signals and allows a better estimation of ssDNA-A488 uptake. Thus, 5 μ*l* of TrypLE Express (commercial trypsin solution) were added to each sample after incubation, and activated at 37ºC for 1 minute. Afterwards, cells were washed and the protocol was resumed as

described in section 4.2.1. TrypLE is a commercial grade trypsin solution and its formulation has not been divulged by the manufacturer. As such, the volume of enzyme solution was chosen after preliminary trials to minimize losses in cell number and viability. The resulting concentration was estimated in the 0.2-1 μM range using concentration values found in other commercial formulations of trypsin. Figure 4.4 depicts the results from this procedure.

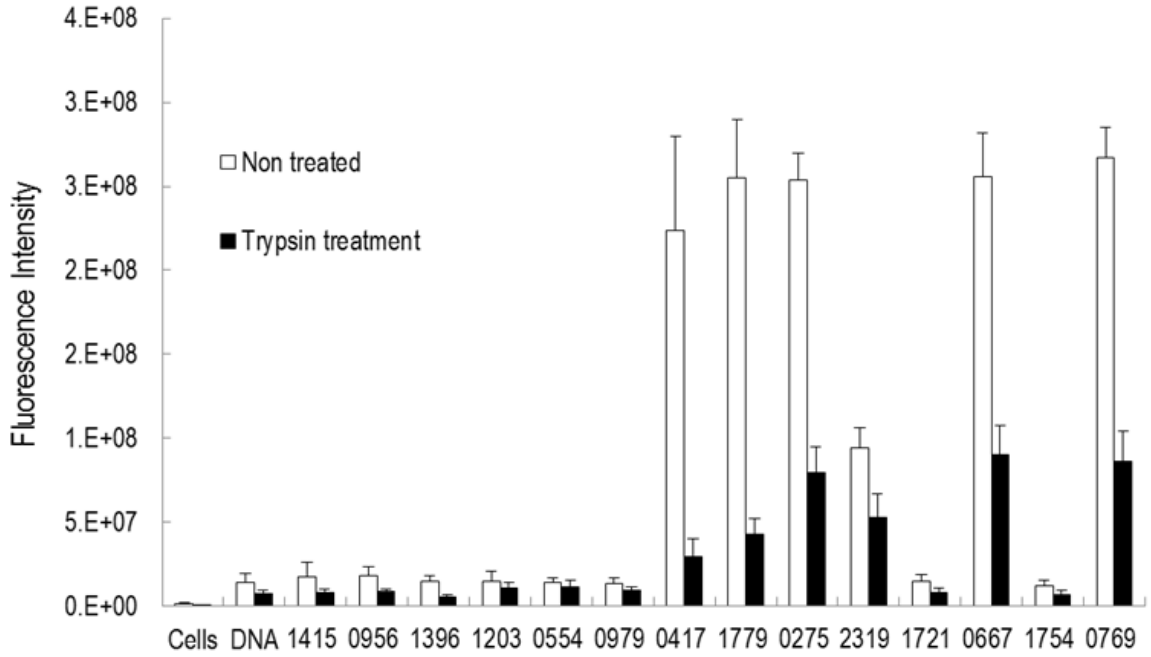


Figure 4.4 – Effect of protease treatment on viralCPP efficiency – Peptide concentration was kept at 5 μM. White bars represent untreated samples (n=5), black bars represent samples treated with 5μl of TrypLE Express (n=3). Error bars are represented as SEM.

From the decrease in fluorescence intensity after addition of trypsin to each sample, it is immediately clear that the protease had a pronounced effect on the amount of probe retained by cells. In peptides belonging to the group of non-CPPs (viralCPPs 1415-0979, 1721 and 1754) as well as in the oligonucleotide control, the estimate of uptake had about 2-fold decrease. It is possible that this effect was related to fragilization of cells by protease activity, perhaps resulting in leakage of intracellular content or in a greater amount of pellet loss by washing. However, in the remaining viralCPPs, reduction in fluorescence intensity, and hence in delivery, was much higher. These findings suggest that a considerable fraction of ssDNA-A488 had been retained in each sample due to membrane binding of viralCPP:ssDNA-A488, and not actually to viralCPP promoted uptake of oligonucleotide cargo. The percentual decrease in fluorescence was particularly large for viralCPPs 0417 and 1779. Conversely, viralCPP 2319 had the smallest percentage of decrease in A-

488 signal. Electrostatic attraction between negative moieties in lipid headgroups, or proteoglycans in the extracellular matrix, and positively charged amino acid residues, has been implicated in CPP activity. ViralCPPs which experienced higher decrease in fluorescence signal were rich in Rs and Ks. In these samples, it is possible that excessive affinity towards membranes had been detrimental to delivery. Thus, a number of viralCPP:ssDNA-A488 may have not been internalized, instead remaining attached to the outer leaflet of the cell surface or inserted in the bilayer. This hypothesis would explain the decrease in fluorescence after enzymatic treatment. ViralCPPs would absorb to membranes and serve as "anchors" for ssDNA-A488 at the surface of cells. These oligonucleotides would therefore be released upon cleavage of the peptides exposed to the extracellular medium by trypsin.

In future trials this experiment may be repeated substituting trypsin with a DNAse solution. This would have the advantage of removing non-internalized probe while minimizing disruptions to cell viability and morphology relative to trypsin.

4.2.3. Live cell imaging of viralCPP:ssDNA-A488 internalization by CM

Microscopy observations are an important procedure to confirm CPP activity because delivery of cargo can be visually confirmed. CM is particularly useful as images are collected from selected focal planes and therefore distinct cell layers can be imaged to assay peptide internalization. Primary CM observations had been conducted with HEK 293 cells (the cell line used in previous experiments), but image quality was hindered by the characteristic over-confluency and chaotic growth of this cell line. Therefore BHK 21 cells were used in substitution to improve the assessment of delivery. Although CPP uptake can vary between cell lines *in vitro*, observations with BHK 21 were consistent with results from FS and FC in HEK 293 (Figure 4.2). This was expected given that both cell types are originated from mammalian kidney. Hoechst probe was used to stain cell nuclei (blue), while ssDNA oligonucleotide molecules were imaged due to A488 labelling (green). Membrane staining was attempted with Cellmask Deep Red, to mark cellular boundaries and co-localize viralCPPs:ssDNA-A488 and cell membranes. However, this probe was found to be too erratic to obtain properly stained samples, perhaps due to membrane metabolism, flip-flop and recycling. As very poor contrast was systematically obtained upon addition of viralCPPs:ssDNA-A488, examples of samples stained with Cellmask Deep Red are presented as supplementary data in Annex B3.

Figure 3.5 shows results obtained from a control of ssDNA-A488 and five peptides. ViralCPP 0554 had been excluded as a CPP and was therefore used a negative control. On the other

hand, viralCPPs 0275, 2319, 0667 and 0769 belong to the group of peptides which displayed CPP activity (Figure 4.4).

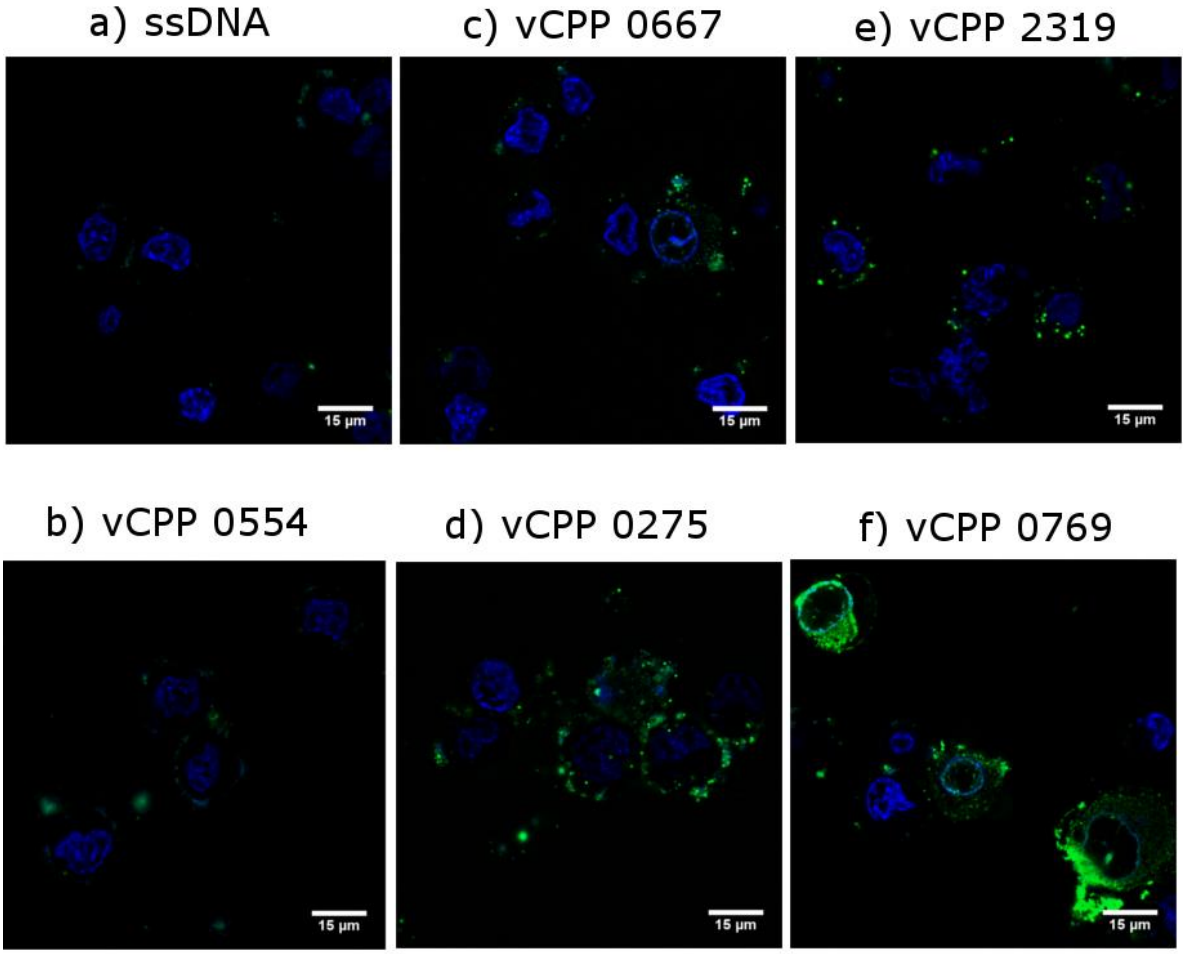


Figure 4.5 – CM of viralCPP-mediated ssDNA Delivery – Composite images of Hoechst staining (Blue channel) and ssDNA-A488 (Green Channel), with 50 µM of peptide and 2 µM of ssDNA. a) ssDNA control; b) viralCPP 0554; c) viralCPP 0667; d) viralCPP 0275; e) viralCPP 2319; f) viralCPP 0769

The top-left image (panel a) shows a field of observation where ssDNA-A488 signal was negligible illustrating that ssDNA-A488 alone could not be retained by cells. Similarly, in panel b corresponding to viralCPP 0554, small amounts of green fluorescence were obtained in the periphery of a few cells, but no evidence for internalization was detected. On the contrary, the remaining four panels (c-f) exhibited an increase in the fluorescence of A-488 and provided evidence for successful viralCPP-mediated uptake of probe. In this regard, viralCPP 0667 (panel c) was the least efficient CPP, producing a diffuse punctuate pattern. ViralCPP 0275 (panel d) produced a similar but brighter pattern. The most efficient CPPs are shown in right-ward panels on

Figure 4.5, namely viralCPP 2319 (panel e) and viralCPP 0769 (panel f). ViralCPP 2319 also produced a punctuate pattern although the dots were brighter and substantially larger than in other panels. This pattern was suggestive of endosomal uptake, and proximity to the nuclei further supported that viralCPP:ssDNA-A488 internalization was successful. The last panel, corresponding to viralCPP 0769, shows intense green fluorescence on the inside of cells, confirming uptake of ssDNA-A488. In this case the pattern was more uniform throughout the cytosol instead of dot-like. Due to this fact, it would be interesting to look into the possibility that viralCPP 0769 may enter cells via direct translocation. This could be achieved through the study of internalization kinetics using FC (118).

CM observations were in accordance with findings described in section 4.2.1, and have confirmed viralCPPs 2319 and 0769 as CPPs. For viralCPP 0275 and 0667 uptake also seemed to have taken place, albeit to lower extent. However, these observations can be optimized for a more robust assessment of viralCPP-mediated uptake of cargo. In this regard, membrane staining with a reliable probe and observing the cells through brightfield transmitted light fluorescence microscopy are future objectives. CM can also be useful in determining the internalization mechanism of each viralCPP. This could be achieved by incubating the peptides along with inhibitors and probes for different pathways of endocytosis to confirm or exclude their role in uptake.

4.2.4. ViralCPP cytotoxicity determination through the Live/Dead assay

The cytotoxicity profile of viralCPPs was evaluated with the Live/Dead mammalian cell kit. This is a two colour assay which uses calcein-AM to stain live cells through green fluorescence, and ethidium homodimer-1 to stain dead cells with red fluorescence. HEK 293 cells were incubated with viralCPPs at 5 µM and 10 µM, stained with the Live/Dead kit and analyzed by FC.

Figure 4.6 provides viralCPP cytotoxicity profiles in terms of cell death (upper panel) and alterations to morphology (lower panel). Negative controls consisted of live cells, while exposure to 70% isopropanol was used as reference for cell death. In general all peptides were well-tolerated by HEK 293 cells, with values inferior to 5% regarding cell death and to 10% in damage. PepM and viralCPP 2319 were exceptions to this pattern. Both peptides displayed similar profiles in either panel, exhibiting considerable toxicity at 10µM. However at half that concentration, cytotoxicity was diminished. In fact, pepM's toxicity dropped to less than 10% in cell death and counts of cells with damaged morphology. However, viralCPP 2319 retained a high (20%) propensity for inducing cell damage. As previously mentioned, a decrease in viralCPP concentration from 10 µM to 5 µM

did not overly alter potency, but these results show that it does alleviate toxicity. Therefore using the peptides at 5 μM would signify an improvement in therapeutic index.

The results in Figure 4.5 suggest that viralCPPs do not exhibit levels of toxicity capable of impairing their development as DDSs. However cytotoxicity tests *in vitro* are a very preliminary form of screening. Extensive assaying *in vivo* would be required to establish a safety profile and optimize the systems before any form of clinical translation could be envisioned for these peptides.

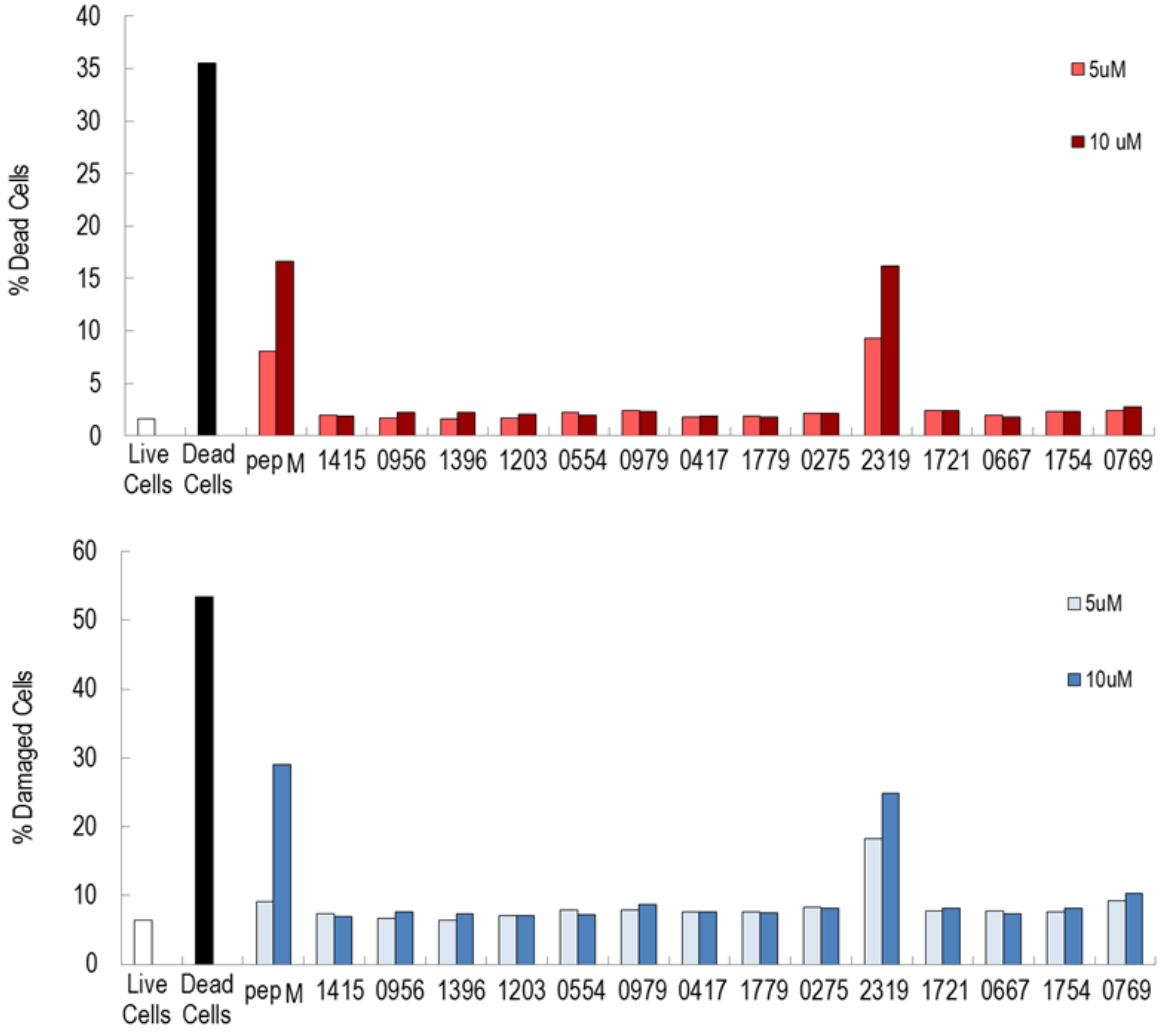


Figure 4.6 – viralCPP toxicity assessment by FC Life/dead assay – Controls of live and dead cells are represented in white and black respectively. (Upper panel) Percentage of cells stained with dead marker (ethidium homodimer-1), peptide concentrations are represented in shades of red. (Lower panel) Percentage of cells with morphology outside of healthy gate, peptide concentrations are represented in shades of blue.

4.3. Biophysical assessment of viralCPPs

The third stage of this work consisted in evaluating the physical-chemical properties of viralCPPs and conjugates of viralCPP and ssDNA. The objective of these tests was to identify possible reasons why some viralCPPs acted as DDSs, while others did not. Accordingly, the affinity of each viralCPP for model membranes was tested via membrane dipole potential sensing with the probe di-8-ANEPPS. Afterwards, formation of secondary structures in aqueous and lipid environments was investigated by CD. In the final place, DLS was used to evaluate the capacity of each peptide to form stable aggregates with ssDNA cargo.

4.3.1. Membrane affinity assayed by di-8-ANEPPS dipole potential sensing

Macromolecules interacting with lipid bilayers alter their membrane dipole potential. As such, dipole potential sensing can be exploited to assess the affinity of MAPs for lipid membranes using voltage-sensitive fluorescent probes such as di-8-ANEPPS (141). For this end, a suspension of 200 μM POPC LUVs, as well as of 200 μM POPC/POPS (4:1) LUVs, were both prepared containing 0.5 % (n/n) di-8-ANEPPS. These vesicles were used to determine whether each viralCPP displayed an affinity towards model membranes at 10 μM, the maximum concentration used in previous assays. The plots shown in Figures 4.7, 4.8 as well as B.4.3 and B.4.4 (in Annex B.4) are differential spectra of di-8-ANEPPS in which peak amplitude relays the magnitude of peptide-membrane interactions. Figure 4.7 depicts differential spectra from peptides representing two groups of viralCPPs. The upper panels of this image belong to viralCPP 1415 representing the group of peptides without affinity for model membranes. Lower panels of the figure display viralCPP 0769 which could extensively interact with lipid vesicles and therefore represents the group of peptides with membrane affinity. Differential spectra belonging to the remaining viralCPPs are presented in Annex B.4 on Figure B.4.3.

The upper panels of Figure 4.7 demonstrate that viralCPP 1415 could not interact with POPC or POPC/POPS LUVS as its differential spectra did not assume a sinusoidal behavior in either case. Conversely, for viralCPP 0769 extensive interactions were reported by the amplitude of the sinusoidal curve determined in POPC/POPS LUVS, and to a lower extent in POPC LUVs. From the spectra present in Figure B.4.3 it was observed that no viralCPP apart from viralCPP 0769 displayed an affinity for zwitterionic POPC membranes. Moreover, viralCPPs 1415, 0956, 1396, 1203, 0979, 0417, 1779, 1721 and 1754 could not interact with POPC/POPS LUVs. Conversely, viralCPPs 0554, 0275, 2319, 0667 and 0769 displayed an affinity for the anionic POPC/POPS

vesicles. A higher magnitude of interaction between viralCPPs and POPC/POPS model membranes was expected relative to POPC LUVS. This is due to the fact that net negative charge (conferred by POPS) is thought to aid in MAP activity by potentiating electrostatic attraction to the cationic residues of the peptides.

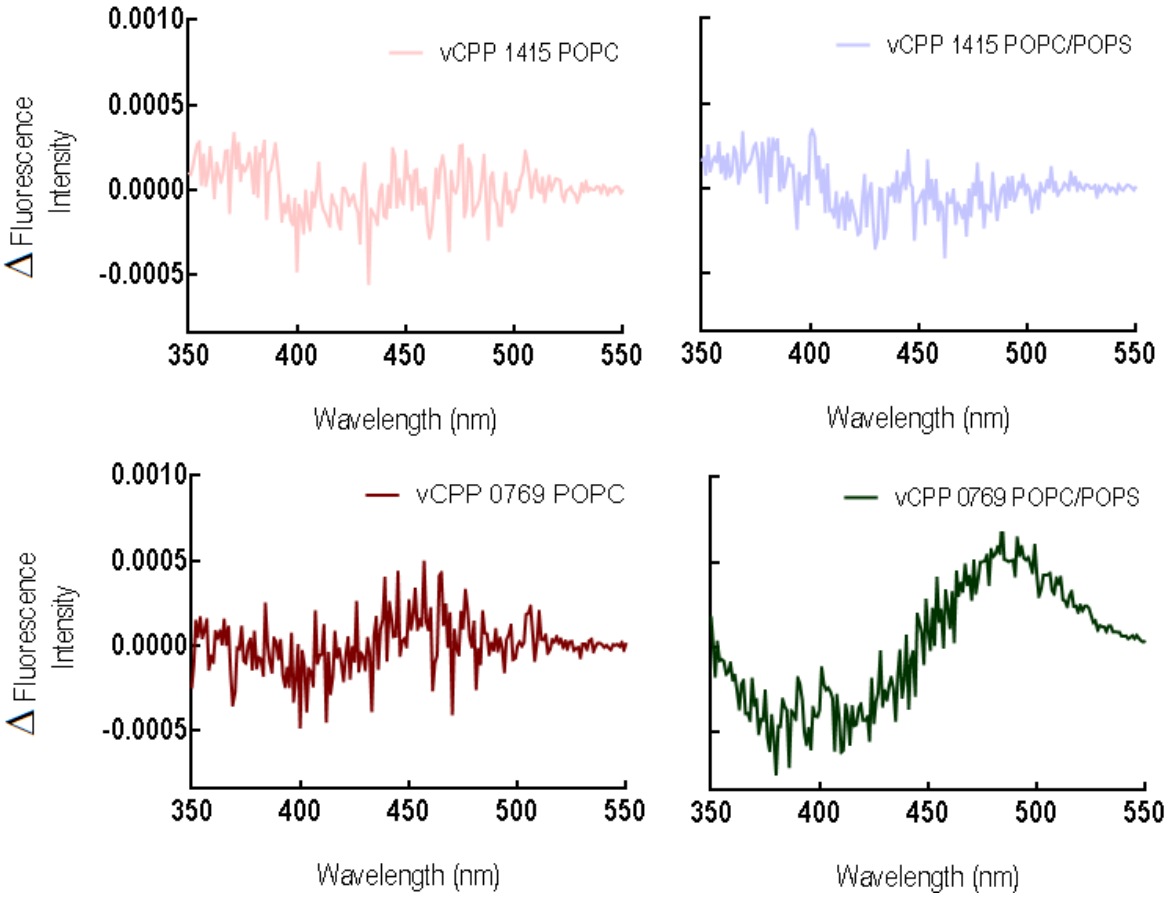


Figure 4.7 – Di-8-ANEPPS dipole potential sensing of viralCPP-membrane interactions – Concentration of potentiometric probe was kept at 0.5 % (n/n), and viralCPPs were tested at 10 µM (Left panels) Differential spectra of interactions reported by di-8-ANEPPS in 200 µM POPC LUVs for viralCPP 1415 (pink) and viralCPP 0769 (brown) (Right panels) Differential spectra of interactions reported by di-8-ANEPPS in 200 µM POPC:POPS (4:1) LUVs for viralCPP 1415 (blue) and 0769 (green).

Because viralCPPs 0554, 0275, 2319, 0667 and 0769 were shown to have an affinity for model membranes, these were assayed at incremental concentrations (1, 5 and 10 µM), and assessed through dual wavelength ratiometric measurements (137). The left panel on Figure 4.8 shows an example of the differential spectra obtained for the three concentrations of viralCPP 0769. Plots belonging to the remaining peptides can be consulted in Figure B.4.4. ViralCPP 0275 was excluded from this assay because no evidence for interaction was found at concentrations lower than 10 µM.

To quantify the magnitude of interaction, the ratiometric approach consists in calculating **R**, the ratio of fluorescence intensities at wavelengths where the crests of the sinusoidal curves occur (e.g. $\mathbf{R}_{\text{viralCPP 0769}} = I_{385}/I_{500}$). **R** was normalized by $\mathbf{R}_0$, the ratio of the intensities for the spectrum of di-8-ANEPPS in undisturbed LUVs (absence of viralCPPs). These results are illustrated in Figure 4.8's right panel.

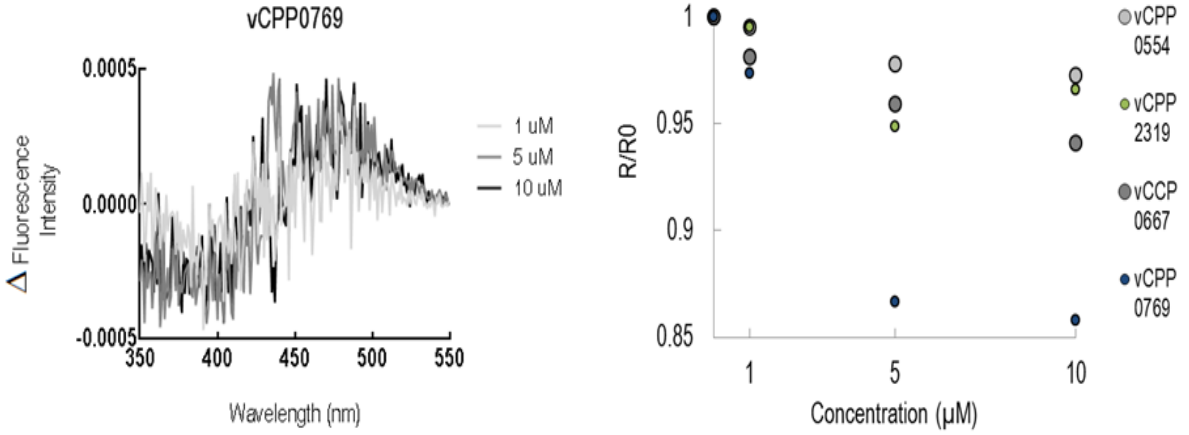


Figure 4.8 – viralCPP dual wavelength ratiometric measurements – (Left panel) Illustration of the shift in amplitude of viralCPP 0769 differential spectra due to increase in concentration. (Right panel) Relative magnitude of viralCPP-POPC/POPS LUV interactions quantified via normalized ratios of fluorescence intensity, mean values from three replicates.

According to these results, viralCPP 0769 induced the highest membrane perturbation of the viralCPP set (higher decrease in $\mathbf{R}/\mathbf{R}_0$) followed by viralCPP 2319, viralCPP 0667 and viralCPP 0554. All of these peptides caused a shift of the excitation spectra of di-8-ANEPPS towards higher wavelengths (Figure B.4.2), *i.e.* a red-shift. This is confirmed by their sinusoidal differential curves (Figure B.4.4), which had negative peaks at lower wavelengths, and positive peaks at higher values (the symmetrical happens for blue-shifts, resulting in increments of **R** with increasing membrane perturbation). Red-shifts therefore relay that these viralCPPs decrease membrane dipole potential as a result of their interactions with POPC/POPS LUVs. The highest magnitude of POPC/POPS LUV membrane dipole potential reduction was caused by viralCPP 0769. It is interesting to note that in CM images (Figure 4.5, panel f) viralCPP 0769-mediated uptake displayed a spread of ssDNA-A488 signal throughout the cytoplasm. This may be explained if the affinity of the peptide for membranes is indicative of direct translocation or of the capacity for endosomal disruption. However, such a hypothesis requires appropriate validation with biophysical (e.g. evaluate the

extent of partition into membranes) and cellular (e.g. FC uptake kinetics) approaches in future studies.

The sequence of viralCPP 0769 is enriched in hydrophobic leucines, with stretches of R residues at both termini. This configuration might confer amphipathic properties to the peptide which are translated into an affinity for membranes and consequent CPP activity. On the other hand, viralCPPs 2319 and 0667 are both highly cationic peptides and thus expected to electrostatically bind to anionic POPC/POPS LUVs. ViralCPP 2319 has three tryptophans and an isoleucine residue, which are hydrophobic moieties that should aid in membrane insertion, but viralCPP 0667 has little content in hydrophobic amino acid residues. However, despite these differences, viralCPP 2319 and 0667 exhibited similar degree of interaction. Conversely, viralCPP 0554 exhibits a much lower magnitude of positive net charge but has higher percentage of hydrophobic amino acids in its sequence. Although excluded as a CPP, this peptide was still found to have an affinity for POPC/POPS vesicles. This data suggests that peptide-membrane interactions and CPP activity depend on a compromise between electrostatic and hydrophobic interactions, and that an excess of one of these factors might actually be detrimental to binding and/or insertion (51, 157). Such findings may explain the behavior of the remaining peptides. In particular, viralCPPs 0417 and 1779 are highly cationic but did not interfere with the physical environment of membrane models, perhaps due to reduced hydrophobicity. Another hypothesis is that they could not be correctly assayed via dipole potential sensing due to the promotion of extensive membrane aggregation. Furthermore, LUVs are membrane models, and although useful in unveiling the biophysical determinants of MAPs, cannot fully replicate interactions with living cell membranes. Therefore, the results described in this section do not necessarily contradict previous data suggesting that viralCPPs 0417 and 1779 have an affinity for biological membranes (Figures 4.2 and 4.4). Furthermore the peptides had previously been assessed in conjugation with ssDNA molecules, and it is reported in the literature that cargo can alter CPP behavior (59, 117). As such, a future set of experiments will assess the affinity of viralCPP and ssDNA conjugates for model membranes.

It should be noted that the concentration of di-8-ANEPPS was kept to a minimum (1 molecule per 200 lipids) so as not to alter the morphology of the model membranes. Peptide concentrations were equally maintained in the range of previous assays (up to 10 μM). Therefore, to improve signal/noise ratio and the sensitivity of the test, higher concentrations of probe and/or viralCPPs could also be tested. Other assays can be conducted to extract additional information on viralCPP-membrane interactions. For instance, partition coefficients of peptide and peptide/cargo

complexes can be determined by spectroscopy (137) or zeta-potential (117) measurements, and binding affinity could be estimated by surface plasmon resonance (160).

4.3.2. Determination of secondary structures with CD

Folding into typical secondary structures, such as α-helices and β-sheets, is postulated as an important biochemical determinant for MAP activity (157). Indeed, secondary amphipathicity acquired by conformational changes seems to play a role in peptide binding and/or disruption of lipid bilayers (67). However, as previously mentioned, the impact of this biochemical trait on CPP activity remains controversial (40, 42-44). To assess the formation of secondary structures through CD, a set of viralCPPs was first selected according to PSIPRED and MCPep helical content predictions (Annex A). PSIPRED (150) is a server for analysis of protein sequences, which contains algorithms for prediction of peptide secondary structure. MCPep (153) is an online platform for Monte Carlo simulation of MAP models in membrane and aqueous environments. ViralCPPs 1721, 0275, 2319 and 0769 were selected for CD measurements because they were predicted to have the highest propensity for acquiring secondary structures. ViralCPP 1721 is 1 of the 8 peptides excluded as CPPs, contrary to viralCPPs 0275, 2319 and 0769 which were efficient in the delivery of ssDNA into HEK cells (Figure 4.2).

Initially, POPC/POPS (4:1) LUVs were employed in CD measurements to maintain the same conditions as past experiments. Incremental LUV concentrations (up to 6 mM) were added to 100μM of viralCPP 2319, as indicated in Figure 4.9 and Table 4.3. Spectra were collected between 200 nm and 240 nm (substantial noise was observed at wavelengths bellow 200 nm due to LUV scattering and aggregation) and the data was plotted as mean residue ellipticity. As Figure 4.9 demonstrates, viralCPP 2319 spectra obtained with POPC/POPS LUVs did not correspond to any of the well-defined CD patterns characteristic of secondary structures (Figure 2.5), suggesting that the peptides mainly remain as random coils. However a shift towards a signature more suggestive of helical features can be observed with an increase in LUV concentration. To quantify the percentage of α-helical content, each spectrum was consequently uploaded into the K2D3 server (155). This tool uses a neural network algorithm to quantify the percentage of secondary structures in CD information uploaded by users by comparison with a reference set of protein CD spectra kept by the server. Results obtained from K2D3 are represented in Table 4.3. This data confirms that α-helix content was incremented in viralCPP 2319 along with the increase in the concentration of lipid vesicles. However, it should be noted that secondary structure prediction with K2D3 does not provide a reliable evaluation of secondary structure content in peptides (142). Therefore, Table 4.3

should be held qualitatively, to support the progressive modification of secondary structures in viralCPP 2319 resulting from the addition of LUVs to the aqueous environment, but not as an accurate estimation of helical content.

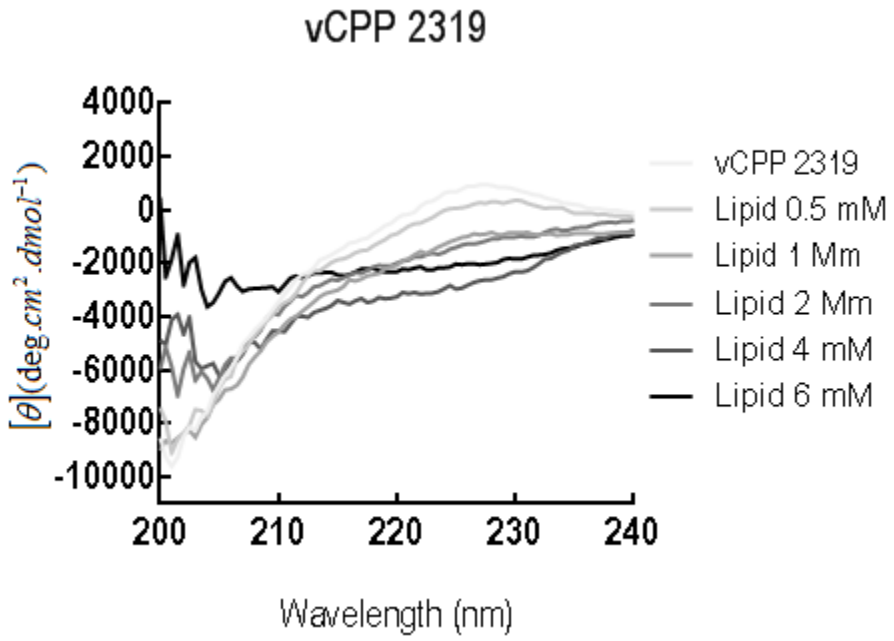


Figure 4.9 – CD spectra of viralCPP 2319 in POPC/POPS (4:1) LUVs – Secondary structure profile of viralCPP 2319 (100 µM) reported via CD in increasing concentrations of POPC/POPS (4:1) LUVs

Table 4.3 – K2D3 prediction of the helical content in CD spectra of viralCPP 2319 in POPC/POPS LUVs

viralCPP	2319					
LUV [POPC] (mM)	0	0.5	1	2	4	6
%α-helix$_{200\text{-}240\ nm}$	90	93	94	94	95	95

Attempts to improve CD signal were conducted using POPC/POPS (4:1) SUVs. SUVs minimize light scattering and are often used to obtain CD measurements with MAPs. Figure 4.10 illustrates CD spectra for viralCPPs 1721, 0275, 2319 and 0769 at 100µM in increasing concentrations of SUVs. Table 4.4 contains the results from the analysis of the spectra by K2D3. CD measurements in SUVs did improve viralCPP 2319's spectrum, as Figure 4.10 illustrates, but still no well-defined structures could be inferred for any of the peptides. Both viralCPP 1721 and 0275 exhibited irregular structures in aqueous environment and upon addition of lipid membranes.

Conversely, for viralCPP 2319 and viralCPP 0769 a shift towards an increase in helical content could be observed upon SUV addition.

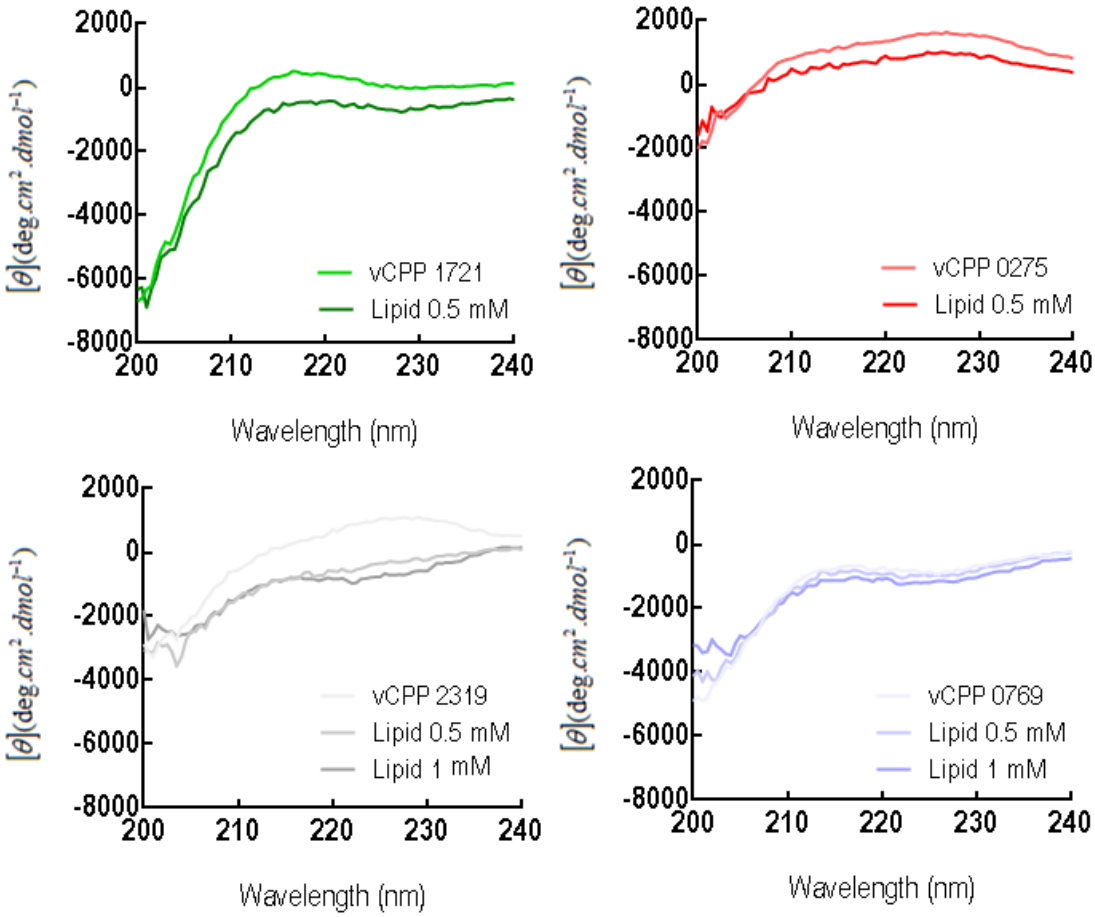


Figure 4.10 - CD spectra of viralCPPs in POPC/POPS (4:1) SUVs - Secondary structure profile of viralCPPs 1721 (green), 0275 (red), 2319 (grey) and 0769 (blue) (100µM) reported via CD in increasing concentrations of POPC/POPS (4:1) SUVs

Table 4.4 – K2D3 prediction of the helical content in CD spectra of viralCPPs in POPC/POPS SUVs

viralCPP	1721		0275		2319			0769		
SUV [POPC] (mM)	0	0.5	0	0.5	0	0.5	1	0	0.5	1
%α-helix$_{200\text{-}240\ nm}$	68	93	1.3	1.7	18	94	94	94	94	95

While adoption of typical conformations is often cited as determinant for membrane activity, cationic CPPs such as Tat and R_9 have been reported to remain unstructured, even in the

presence of zwitterionic or anionic membrane models (42). ViralCPPs studied in this section are also highly cationic, and as peptides derived from a natural source were not tailored to acquire primary or secondary amphipathic properties. Therefore, CD results here described suggest that these viralCPPs may be part of a class of naturally disordered CPPs. These results are also partially concordant with the simulations that were previously mentioned (Annex A). Indeed, MCpep correctly predicted that viralCPP 0769 would have the maximum helical propensity in the set, and that minimal shifts in conformation between aqueous medium and anionic membranes would be expected for viralCPP 1721 and 0275.

4.3.3. ViralCPP/ssDNA aggregation studied by DLS

This section marks the conclusion of the preliminary run of biophysical studies on viralCPPs. Specifically, the formation of electrostatically bound supramolecular complexes between viralCPPs and ssDNA oligonucleotides was studied through DLS. This technique allows the determination of the mean hydrodynamic diameters of nanometer scaled particles in solution. However, the hydrodynamic diameter may only give a reliable estimate of dimension if the particles are spherical. This is not the case for the peptides, nor necessarily for peptide/oligonucleotide conjugates. Nonetheless, DLS can relate the apparent size at which dispersed particles are more stable. Therefore DLS measurements were used to monitor the mean hydrodynamic diameters of peptides in the absence and presence of ssDNA oligonucleotides. Increases in particle size were attributed to the formation of peptide/oligonucleotide complexes by non-covalent binding of ssDNA with viralCPPs. Concentrations of 100 μM and of 20 μM were tested for viralCPPs and ssDNA, respectively, to optimize measurement conditions. Histograms of the distribution in hydrodynamic diameters for each viralCPP are provided in Annex B5. According to this data, viralCPPs often demonstrated bi-disperse profiles, i.e. a population of particles with hydrodynamic diameters around 1-10 nm and another population with much larger diameters. Therefore, for each sample, the two populations were accounted separately by setting a threshold at 10 nm in average hydrodynamic diameter. Particles with hydrodynamic diameters at lesser values than 10 nm were considered well-dispersed, whereas those of larger diameter were held as micro-aggregates. The mean hydrodynamic diameter of each viralCPP population was calculated and is displayed per viralCPP, along with the associated percentage in particle number, in Tables 4.5 and 4.6.

Table 4.5 and Figure B.5.1 contain the results from DLS experiments with viralCPPs excluded as DDSs in section 4.2.1. ViralCPPs 1415, 1396 and 1203 showed very residual or nonexistent aggregation before addition of ssDNA. On the other hand, viralCPPs 0956, 0554, 0979

and 1754 acquired bimodal distributions of hydrodynamic diameter before addition of oligonucleotides. The majority of peptides in each of these samples (60-80%) clustered at a scale of around 1nm, but some aggregation was also present (20-40%). Conversely, viralCPP 1721 was found to be almost completely aggregated in the absence of ssDNA. Upon the addition of oligonucleotides, four of these viralCPPs (1415, 1203, 0979 and 1754) retained a considerable population of particles with a hydrodynamic diameter of less than 10 nm. This finding supports the hypothesis that they did not form stable electrostatic complexes with the ssDNA molecules. In this regard, viralCPP 1415 and 1203 were the least efficient in binding to ssDNA because conjugation was inexistent. ViralCPPs 0979 and 1754 were similarly poor in terms of electrostatic binding to ssDNA oligonucleotides. In this two peptides only about 30% of the total number of particles was shown to form complexes. However these complexes had about half the size (~300 nm) of those formed in the absence of nucleic material (> 600 nm) indicating that conjugation was partly successful. Similarly, viralCPP 1721 was again found to be completely aggregated, but its hydrodynamic diameter also decreased by half in the presence of ssDNA.

Table 4.5 – Hydrodynamic diameter of low efficiency viralCPPs

(μM)	$\bar{d}$(nm)	Low Score viralCPP 1415	0956	1396	1203	0554	0979	1721	1754
[Peptide] (100)	0-10	1.7 93.3%	1.4 63.3%	1.2 100.0%	1.2 100.0%	1.2 86.6%	1.5 66.6%	3 3.3%	1 76.6%
	10-1000	1043 6.7%	481 36.6%	0.0 0.0%	0.0 0.0%	604 13.3%	686 33.3%	426 96.6%	823 23.3%
[Peptide + ssDNA] (100:20)	0-10	1.9 100%	2 30.0%	0.0 0.0%	2.4 93.3%	0.0 0.0%	2.6 76.6%	0.0 0.0%	2.5 73.3%
	10-1000	0.0 0.0%	173 70.0%	408.0 100.0%	1163 6.7%	848.6 100.0%	349 23.3%	236.6 100.0%	323 26.6%

Collectively, these results suggest that for viralCPPs 1415, 1203, 0979, 1754 and 1721, poor binding to ssDNA molecules may have been an important justification for their deficiency in intracellular delivery of the oligonucleotides (Figure 4.2).

Although having the lowest magnitude of positive charges (+1) in the set, viralCPP 0956 was still capable of forming complexes with oligonucleotides. Both this peptide and viralCPP 1415 have anionic residues in their sequence. In viralCPP 1415 the effect of the aspartic and glutamic acid residues might explain the complete absence of electrostatic affinity towards nucleic acids, due to repulsion to phosphate groups. Similarly, viralCPPs 1203 and 0554 have the same magnitude of positive net charge (+ 6), but the first peptide could not bind to oligonucleotides, possibly due to the presence of an aspartic acid in its amino acid sequence. Furthermore the sequence of viralCPP 1721 also has two negatively-charged amino acid residues, but this peptide was shown to aggregate even in the absence of ssDNA, and therefore it was not clear if conjugation occurred.

Table 4.6 and Figure B.5.2 contain DLS data referent to those viralCPPs which were found to acts as CPPs in the delivery of ssDNA *in vitro*.

Table 4.6 - Hydrodynamic diameter of high efficiency viralCPPs

(μM)	$\bar{d}$ (nm)	High Score viralCPP 0417	1779	0275	2319	0667	0769
[Peptide] (100)	0-10	1.5 100.0%	1.4 90.0%	1.6 93.3%	1.9 97.8%	1.3 96.7%	0.0 0.0%
	10-1000	0.0 0.0%	605 10.0%	1233 6.7%	14 2.2%	2010 3.3%	163.2 100.0%
[Peptide + ssDNA] (100:20)	0-10	0.0 0.0%	0.0 0.0%	0.0 0.0%	0.0 0.0%	0.0 0.0%	0.0 0.0%
	10-1000	991.6 100.0%	910.7 100.0%	744.7 100.0%	104.8 100.0%	810.0 100.0%	911.9 100.0%

In general, viralCPPs 0417, 1779, 0275, 2319, 0667 and 0769 demonstrated a different behavior than the non-CPPs. For all peptides, except viralCPP 0769, no propensity for spontaneous aggregation before ssDNA addition was observed. Furthermore, all 6 viralCPPs were completely conjugated upon the addition of ssDNA oligonucleotides, forming monomodal distributions of hydrodynamic diameters. These findings are not surprising, given that the peptides have greater magnitude of positive net charge (> + 8), and no anionic amino acid residues in their sequences. Therefore these viralCPPs are more likely to repel each other and remain well-dispersed in solution, while simultaneously demonstrating higher affinity for electrostatic conjugation with negatively-charged ssDNA molecules. Among this group, two particularly interesting cases were those of

viralCPP 2319 and of viralCPP 0769. As Table 4.6 shows, viralCPP 2319 was fully conjugated with ssDNA forming a monodisperse population at 100 nm in hydrodynamic diameter. This may have happened because it had the highest magnitude of positive net charge (+ 15) and was therefore able to form smaller stable conjugates by binding more oligonucleotides per peptide. As such, in less cationic viralCPPs electrostatic "nucleation" would proceed until larger, presumably neutral structures could be formed. Indeed, in this set of 6 viralCPPs, higher amount of charge was correlated (Pearson coefficient of -0.9) with a decrease in hydrodynamic diameter. Conversely, viralCPP 0769 was shown to fully aggregate before nucleic acid addition, but a shift in hydrodynamic diameter could still be observed from 160 nm to 900 nm, which together with previous data is indicative of successful conjugation.

The electrostatic conjugation between viralCPPs and ssDNA oligonucleotides is a pre-requisite for their development as DDSs. However, aggregation profiles are also important in the optimization of delivery conditions. In an article published by Morris *et al*. (161) the authors found that the *in vitro* delivery of cargo mediated by the CPP pep-1 was dependent on the molar ratio between cargo and vector. In this study, molar ratio influenced the size of the pep-1/cargo complexes and consequently the efficiency of delivery. A molar ratio of 20:1 (pep-1:cargo) was found to produce particles with 100 nm in size, as evaluated by DLS and scanning electron microscopy. This population of particles sized at around 100 nm was highly efficient in transfecting cells. In opposition, different ratios of pep-1 to cargo produced larger particles (> 500 nm) which were much less efficient in delivery. These findings by Morris *et al*. are reminiscent of those obtained with viralCPP 2319. Therefore, the fact that viralCPP 2319 was capable of forming conjugates centered around 100 nm in hydrodynamic diameter might partly explain why it was an efficient viralCPP in the delivery of ssDNA oligonucleotides *in vitro*.

4.4. From viralCPPs to viralMAPs

From the set of 14 viralCPPs which have been assayed *in vitro*, 6 (viralCPPs 1779, 0417, 0275, 2319, 0667 and 0769) displayed CPP activity. ViralCPP 2319 and 0769 were particularly efficient vectors, and would be good leads for continuing development. Therefore a success rate of at least 15% could be achieved with the current viralCPP sample. Assuming that this ratio can be reproduced in the original pool of 2400 viralCPPs, this means that a minimum of 350 CPP sequences are expected to have been found, confirming the hypothesis that structural viral proteins are an excellent source of CPPs.

Apart from CPPs, other classes of peptides exist which exert pronounced effects upon lipid membranes. These are collectively known as MAPs (114). One of the most extensively studied class of MAPs is that of AMPs. These peptides are lytic to pathogens by inducing the disruption of their lipid membranes (65). AMPs include peptides capable of antibacterial (119), antiviral (162) and even anticancer effects (163). Like CPPs, AMPs are mostly comprised of short amino acid sequences. The great majority of these peptides exhibit significant portion of basic amino acid residues and the propensity for acquiring amphipathic and/or helical conformations upon membrane binding (164). Likewise, exceptions exist in both cases such as peptides of larger dimensions, those rich in acidic residues, and others of hydrophobic character (165). Their mechanisms of action are also intersected in terms of biophysical processes, such as the formation of pores or other membrane defects (166). Furthermore, understanding of these mechanisms is still very much incomplete in both cases (67). In fact, many peptides have demonstrated concurrent antimicrobial action and mammalian-cell non-lytic internalization/cargo delivery (167). For instance, potent antimicrobial action was discovered in two well-studied CPPs, pVec and transportan 10 (168), as well as in tat, penetratin, model amphipathic peptide and pep-1 (169). Furthermore, Bobone *et al*. (170) have recently underlined how subtle differences can dictate a tendency for either antimicrobial or cell penetrating activities. Symmetrically, AMPs have also been described to show CPP behavior, such as the hLF peptide from human lactoferrin, LL-37 (169, 171).

Both AMPs and CPPs are highly relevant in the context of biotechnology and nanomedicine. AMPs in particular have grown very rapidly as a research field in past few years. This is understandable given the urgency in identifying powerful substitutes for conventional antibiotics due to the grievous scenario of bacterial resistance (172-174). For instance, AMP research abounds in peptides derived from natural proteins found in animals, e.g. insects, reptiles, amphibian, fowls and mammals, to plants and even bacteria (164). However, Figure 4.11shows that the total percentage of AMPs derived from viruses is close to inexistent. As with CPPs, the precise set of properties that condition the interactions between AMPs and biological membranes is not known. However, an increase in the number of AMP databases has allowed the refinement of algorithms to detect AMP propensity in amino acid sequences (130, 175). Table 4.7 shows the great variety in currently existing AMP databases and prediction servers relative to those dedicated to CPPs. Among the tools in Table 4.7, AMPA (176) is a server for the identification of AMP regions in amino acid sequences of proteins. It can therefore be viewed as an analogue of CellPPD, designed for antimicrobial research.

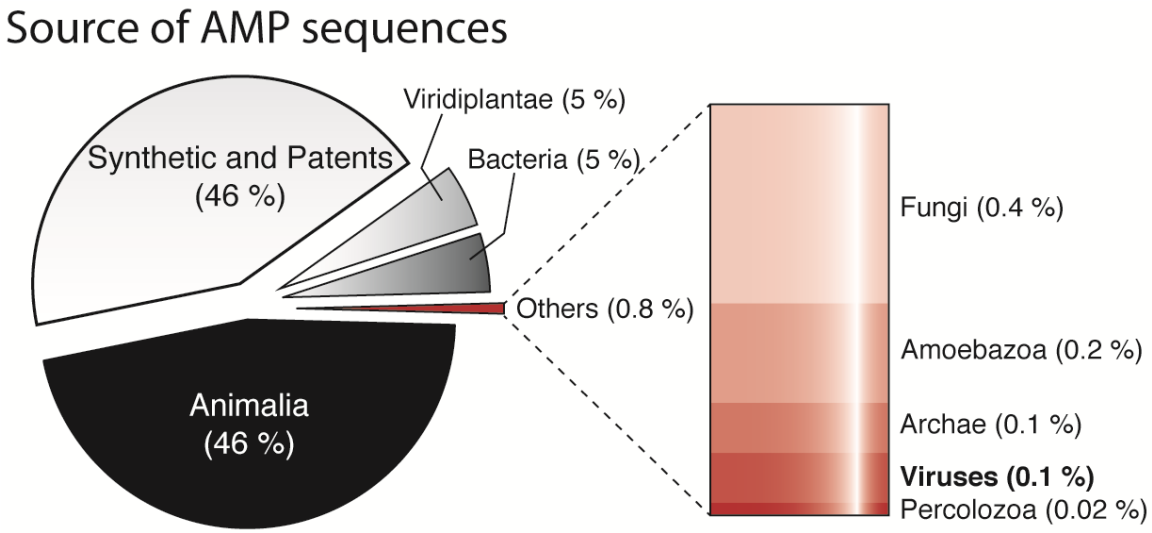


Figure 4.11 - Sources of AMP Sequences – Each percentage was calculated based on the amount of hits from several AMP databases.

Table 4.7 - Currently available databases and prediction tools for AMP and CPP research

	Name	Year	Sequences	Description	Ref
Databases					
AMP					
	YADAMP	2012	2525	Yet Another AMP database	(177)
	APD2	2004	2329	The Antimicrobial Peptide Database 2	(178)
	CAMP	2010	5040	Collection of Antimicrobial Peptides	(179)
	LAMP	2012	5547	Linking Antimicrobial Peptides	(180)
	DAMPD	2011	1232	Dragon Antimicrobial Peptides	(181)
CPP					
	CPPsite	2012	843	Cell-Penetrating Peptide Database	(35)
Prediction Tools					
AMP					
	AMPA	2012		Analyze protein sequence for AMP regions	(182)
	ClassAMP	2012		AMP prediction	(183)
	AntiBP 2	2009		AMP prediction	(184)
	AntiCP	2013		AMP prediction	(185)
	iAMP-2L	2013		AMP prediction	(186)
CPP					
	CellPPD	2013		Analyze protein sequence for CPP regions	(126)
	CPPpred	2013		CPP prediction	(131)

In this work, structural viral proteins were confirmed as a reliable source of CPPs. Given the fact that AMPs are biochemically identical to CPPs and the scenario illustrated by Figure 4.11, it would be highly relevant to search for novel AMP sequences in structural proteins from viruses. Therefore, AMPA will be used to locate novel AMP sequences, or viralAMPs, in the same pool of 270 proteins in which viralCPPs were identified. Thus, the success found with viralCPPs has motivated the prospect of expanding the opportunities for MAP research in the underexplored but highly promising source of structural proteins from viruses.

5. Conclusions and Future Prospects

This work concerned the identification of a type of MAPs, CPPs, from structural proteins of distinct virus families. Consequently, its hypothesis was that these components from the global proteome of viruses are an abundant and relatively unexplored source of peptide sequences which have been tailored by nature to interact with lipid membranes. ViralCPPs tested in this project are clear examples that confirmed this statement. Tat, the primordial CPP, is part of a viral protein, and a few other CPPs have been found in amino acid sequences from elements of capsids and envelopes. As such, the work herein described can be seen as a "back to basics" approach to the overall research linked to MAPs.

In the first stage of this work sequences of CPP propensity were identified in structural viral proteins, through the use of the CellPPD server. According to the obtained results, the machine learning algorithm implemented in the server could not only correctly identify protein sequences of CPP potential, but also predict the expected degree of efficiency of every CPP candidate. This is a very useful feature for peptide research, enabling rational grafting and tuning of natural amino acid sequences. With this application, a pool of thousands of CPP leads was created, and was subsequently refined by clustering and manual selection to produce the final 14 viralCPP set. Selection criteria included biochemical diversity and novelty, in detriment of simply searching for the highest scoring leads.

To test these molecules, a screening protocol was developed. Its advantages included low biological material requirements, simplicity, and rapid assessment of CPP activity by complementary measurements (FS, FC and PR). Probing cargo instead of the peptides themselves is advantageous in that no artifacts stem from hydrophobic moieties introduced in the amino acid sequences, but it implies that cargoes be labeled with fluorophores prior to testing. Through this methodology, it was found that 6 of the 14 peptides had CPP potential to deliver genetic material into cells. Protease treatment and CM analyses provided further evidence that two of the peptides, namely viralCPP 2319 and viralCPP 0769 were the best and more robust candidates for further development as CPPs. For viralCPPs 0417, 1779, 0275 and 0667 further tests have to be conducted in order to settle their potential as CPPs. Additionally; it was found that viralCPPs exhibited low level of toxicity, as assessed by Life/Dead assay, signifying that they could be further optimized as DDSs.

The last stage of this project consisted in preliminary biophysical characterization of viralCPP interactions with lipid membrane models and molecular cargo, namely through di-8-

ANEPPS membrane dipole potential sensing, CD and DLS spectroscopy. In this regard, it was found that poor CPPs were incapable of interacting with membrane models, and in many cases could not be conjugated to ssDNA. On the other hand, high-scoring highly-charged cationic peptides ubiquitously formed complexes with cargo, and viralCPP 2319 and 0769, in particular, were found to induce the highest degree of perturbation in POPC/POPS LUVs. Furthermore, these viralCPPs did not acquire a CD profile related to typical secondary structure conformations signifying that they are part of the cationic CPP family and remain unstructured in lipids. Thus, viralCPPs 2319 and 0769 are the two main candidates for CPP development, whereas viralCPP 0417, 1779, 0667 and 0275 are contenders which must be subjected to further testing.

Future perspectives for the conclusion of this project may include a more in depth study of some of these peptides. Further FS, FC and CM assays with additional cell lines/cargoes can be conducted, incubating cells under endocytosis blocking conditions (low temperature and endocytic inhibitors), determining peptide uptake kinetics, and attempting transfection with a reporter plasmid/gene knockdown approach to ascertain endosomal escape and functional cargo delivery. Other biophysical trials can also be envisioned, such as membrane fusion assays or the determination of partition coefficients through zeta-potential or FS measurements. Following stages of this work will also encompass the identification of viralAMPs in the pool of 270 structural viral proteins. As with viralCPPs, a set of these antimicrobial agents will be validated *in vitro* by employing standard techniques such as determination of minimum inhibitory concentrations, hemolysis and bacterial Live/Dead assays. Finally, bioinformatics approaches might also be undertaken to search for evidence of evolutionary biases of viral structural proteins for MAP content, in order to establish that these regions serve as biochemical domains for interaction with biological bilayers.

The work described in this document bridged multiple fields, such as drug delivery and virology, applying several techniques from diverse areas including bioinformatics, biophysics, and molecular biology. Indeed, MAPs are a recent concept with promising applications in intersecting areas of biomedical relevance. Thus, expanding the pool of promising leads and the continuing efforts to unearth which biochemical features are determinant in this regard are hot-topics in peptide research that can contribute with exciting applications to a wide range of medical problems.

Viruses were one of the primary sources of CPPs, but remain relatively unexplored as a natural reservoir of MAPs. Structural viral proteins have been shaped by evolution to counter the imposing frontier of cell membranes and are therefore a natural reservoir for MAPs that can now be mined through a combination of computational tools and experimental techniques. As such, this

work and its future stages will re-establish viruses, complex molecular machines, as frontrunners in a growing field of biomedical relevance, MAP research.

6. References

1. Zhou J, Rossi JJ. The Potential and Current Progress of Internalizing Molecules in Targeted Drug Delivery. In: Prokop A, editor. Intracellular Delivery - Fundamentals and Applications: Springer; 2011. p. 97-123.
2. Parveen S, Misra R, Sahoo SK. Nanoparticles: a boon to drug delivery, therapeutics, diagnostics and imaging. Nanomedicine : nanotechnology, biology, and medicine. 2012;8(2):147-66. Epub 2011/06/28.
3. Murday JS, Siegel RW, Stein J, Wright JF. Translational nanomedicine: status assessment and opportunities. Nanomedicine : nanotechnology, biology, and medicine. 2009;5(3):251-73. Epub 2009/06/23.
4. Allen TM, Cullis PR. Drug delivery systems: entering the mainstream. Science. 2004;303(5665):1818-22. Epub 2004/03/20.
5. Ibraheem D, Elaissari A, Fessi H. Gene therapy and DNA delivery systems. International journal of pharmaceutics. 2014;459(1-2):70-83. Epub 2013/11/30.
6. Ferrati S, Streiff AK, Srinivasan S, Alexander JF, Bhargava N, Peters AM, et al. Mass Transport via Cellular Barriers and Endocytosis In: Prokop A, editor. Intracellular Delivery - Fundamentals and Applications: Springer; 2011. p. 3-55.
7. Milletti F. Cell-penetrating peptides: classes, origin, and current landscape. Drug discovery today. 2012;17(15-16):850-60. Epub 2012/04/03.
8. Vizirianakis IS, Fatouros DG. Personalized nanomedicine: paving the way to the practical clinical utility of genomics and nanotechnology advancements. Advanced drug delivery reviews. 2012;64(13):1359-62. Epub 2012/09/18.
9. Zhang XQ, Xu X, Bertrand N, Pridgen E, Swami A, Farokhzad OC. Interactions of nanomaterials and biological systems: Implications to personalized nanomedicine. Advanced drug delivery reviews. 2012;64(13):1363-84. Epub 2012/08/25.
10. Giacca M, Zacchigna S. Virus-mediated gene delivery for human gene therapy. Journal of controlled release : official journal of the Controlled Release Society. 2012;161(2):377-88. Epub 2012/04/21.
11. Beija M, Salvayre R, Lauth-de Viguerie N, Marty JD. Colloidal systems for drug delivery: from design to therapy. Trends in biotechnology. 2012;30(9):485-96. Epub 2012/06/08.
12. Agency EM. European Medicines Agency recommends first gene therapy for approval - Glybera offers new medical treatment for patients with severe or multiple pancreatitis attacks due to lipoprotein lipase deficiency 2012 [January 2013]; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2012/07/WC500130146.pdf.
13. Yildiz I, Shukla S, Steinmetz NF. Applications of viral nanoparticles in medicine. Current opinion in biotechnology. 2011;22(6):901-8. Epub 2011/05/20.
14. Barratt G. Delivery to Intracellular Targets by Nanosized Particles. In: Prokop A, editor. Intracellular Delivery - Fundamentals and Applications: Springer; 2011. p. 73-95.
15. Varkouhi AK, Scholte M, Storm G, Haisma HJ. Endosomal escape pathways for delivery of biologicals. Journal of controlled release : official journal of the Controlled Release Society. 2011;151(3):220-8. Epub 2010/11/17.
16. Deshayes S, Morris M, Heitz F, Divita G. Delivery of proteins and nucleic acids using a non-covalent peptide-based strategy. Advanced drug delivery reviews. 2008;60(4-5):537-47. Epub 2007/11/27.

17. Koren E, Torchilin VP. Cell-penetrating peptides: breaking through to the other side. Trends Mol Med. 2012;18(7):385-93.
18. Huang Y, Jiang Y, Wang H, Wang J, Shin MC, Byun Y, et al. Curb challenges of the "Trojan Horse" approach: Smart strategies in achieving effective yet safe cell-penetrating peptide-based drug delivery. Advanced drug delivery reviews. 2013. Epub 2013/02/02.
19. Fonseca SB, Pereira MP, Kelley SO. Recent advances in the use of cell-penetrating peptides for medical and biological applications. Advanced drug delivery reviews. 2009;61(11):953-64.
20. Wang F, Wang Y, Zhang X, Zhang W, Guo S, Jin F. Recent progress of cell-penetrating peptides as new carriers for intracellular cargo delivery. Journal of controlled release : official journal of the Controlled Release Society. 2013;174C:126-36. Epub 2013/12/03.
21. de Figueiredo IR, Freire JM, Flores L, Veiga AS, Castanho MA. Cell-penetrating peptides: A tool for effective delivery in gene-targeted therapies. IUBMB life. 2014. Epub 2014/03/25.
22. Craik DJ, Fairlie DP, Liras S, Price D. The Future of Peptide-based Drugs. Chem Biol Drug Des. 2013;81(1):136-47.
23. Lax R. The Future of Peptide Development in the Pharmaceutical Industry. PharManufacturing: The International Peptide Review. 2010.
24. Frankel AD, Pabo CO. Cellular uptake of the tat protein from human immunodeficiency virus. Cell. 1988;55(6):1189-93. Epub 1988/12/23.
25. Derossi D, Joliot AH, Chassaing G, Prochiantz A. The third helix of the Antennapedia homeodomain translocates through biological membranes. The Journal of biological chemistry. 1994;269(14):10444-50. Epub 1994/04/08.
26. Vives E, Brodin P, Lebleu B. A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. The Journal of biological chemistry. 1997;272(25):16010-7.
27. Schwarze SR, Ho A, Vocero-Akbani A, Dowdy SF. In vivo protein transduction: Delivery of a biologically active protein into the mouse. Science. 1999;285(5433):1569-72.
28. Lee SH, Castagner B, Leroux JC. Is there a future for cell-penetrating peptides in oligonucleotide delivery? Eur J Pharm Biopharm. 2013;85(1):5-11.
29. Pooga M, Soomets U, Hallbrink M, Valkna A, Saar K, Rezaei K, et al. Cell penetrating PNA constructs regulate galanin receptor levels and modify pain transmission in vivo. Nat Biotechnol. 1998;16(9):857-61.
30. Jones AT, Sayers EJ. Cell entry of cell penetrating peptides: tales of tails wagging dogs. Journal of controlled release : official journal of the Controlled Release Society. 2012;161(2):582-91. Epub 2012/04/21.
31. Simeoni F, Morris MC, Heitz F, Divita G. Insight into the mechanism of the peptide-based gene delivery system MPG: implications for delivery of siRNA into mammalian cells. Nucleic Acids Res. 2003;31(11):2717-24.
32. Morris MC, Depollier J, Mery J, Heitz F, Divita G. A peptide carrier for the delivery of biologically active proteins into mammalian cells. Nat Biotechnol. 2001;19(12):1173-6.
33. Futaki S, Suzuki T, Ohashi W, Yagami T, Tanaka S, Ueda K, et al. Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. The Journal of biological chemistry. 2001;276(8):5836-40. Epub 2000/11/21.
34. Wender PA, Mitchell DJ, Pattabiraman K, Pelkey ET, Steinman L, Rothbard JB. The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(24):13003-8. Epub 2000/11/23.

35. Gautam A, Singh H, Tyagi A, Chaudhary K, Kumar R, Kapoor P, et al. CPPsite: a curated database of cell penetrating peptides. Database : the journal of biological databases and curation. 2012;2012:bas015. Epub 2012/03/10.
36. Heitz F, Morris MC, Divita G. Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics. British journal of pharmacology. 2009;157(2):195-206. Epub 2009/03/25.
37. Stalmans S, Wynendaele E, Bracke N, Gevaert B, D'Hondt M, Peremans K, et al. Chemical-functional diversity in cell-penetrating peptides. Plos One. 2013;8(8):e71752. Epub 2013/08/21.
38. Martin I, Teixido M, Giralt E. Design, synthesis and characterization of a new anionic cell-penetrating peptide: SAP(E). Chembiochem : a European journal of chemical biology. 2011;12(6):896-903. Epub 2011/03/03.
39. Schmidt N, Mishra A, Lai GH, Wong GC. Arginine-rich cell-penetrating peptides. FEBS letters. 2010;584(9):1806-13. Epub 2009/11/21.
40. Bechara C, Sagan S. Cell-penetrating peptides: 20 years later, where do we stand? FEBS letters. 2013;587(12):1693-702. Epub 2013/05/15.
41. Elmquist A, Hansen M, Langel U. Structure-activity relationship study of the cell-penetrating peptide pVEC. Biochimica et biophysica acta. 2006;1758(6):721-9. Epub 2006/07/01.
42. Eiriksdottir E, Konate K, Langel U, Divita G, Deshayes S. Secondary structure of cell-penetrating peptides controls membrane interaction and insertion. Biochimica et biophysica acta. 2010;1798(6):1119-28. Epub 2010/03/11.
43. Alves ID, Jiao CY, Aubry S, Aussedat B, Burlina F, Chassaing G, et al. Cell biology meets biophysics to unveil the different mechanisms of penetratin internalization in cells. Biochimica et biophysica acta. 2010;1798(12):2231-9. Epub 2010/02/16.
44. Alves ID, Rodriguez N, Cribier S, Sagan S. Membrane Crossover by Cell-Penetrating Peptides: Kinetics and Mechanisms – From Model to Cell Membrane Perturbation by Permeant Peptides. In: Prokop A, editor. Intracellular Delivery - Fundamentals and Applications: Springer; 2011. p. 179-96.
45. Torcato IM, Huang YH, Franquelim HG, Gaspar D, Craik DJ, Castanho MA, et al. Design and characterization of novel antimicrobial peptides, R-BP100 and RW-BP100, with activity against Gram-negative and Gram-positive bacteria. Biochimica et biophysica acta. 2013;1828(3):944-55. Epub 2012/12/19.
46. Wadhwani P, Strandberg E, van den Berg J, Mink C, Burck J, Ciriello RA, et al. Dynamical structure of the short multifunctional peptide BP100 in membranes. Biochimica et biophysica acta. 2014;1838(3):940-9. Epub 2013/11/13.
47. Alves CS, Kairys V, Castanho MA, Fernandes MX. Interaction of antimicrobial peptides, BP100 and pepR, with model membrane systems as explored by Brownian dynamics simulations on a coarse-grained model. Biopolymers. 2012;98(4):294-312. Epub 2012/11/30.
48. Raagel H, Saalik P, Pooga M. Peptide-mediated protein delivery-which pathways are penetrable? Biochimica et biophysica acta. 2010;1798(12):2240-8. Epub 2010/02/23.
49. Deshayes S, Decaffmeyer M, Brasseur R, Thomas A. Structural polymorphism of two CPP: An important parameter of activity. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2008;1778(5):1197-205.
50. Su Y, Hong M. Conformational disorder of membrane peptides investigated from solid-state NMR line widths and line shapes. J Phys Chem B. 2011;115(36):10758-67. Epub 2011/08/03.
51. Walrant A, Correia I, Jiao CY, Lequin O, Bent EH, Goasdoue N, et al. Different membrane behaviour and cellular uptake of three basic arginine-rich peptides. Biochimica et biophysica acta. 2011;1808(1):382-93. Epub 2010/10/06.

52. Madani F, Lindberg S, Langel U, Futaki S, Graslund A. Mechanisms of cellular uptake of cell-penetrating peptides. J Biophys. 2011;2011:414729. Epub 2011/06/21.
53. Munyendo W, Ly H, Benza-Ingoula H, Baraza LD, Zhou J. Cell Penetrating Peptides in the Delivery of Biopharmaceuticals. Biomolecules. 2012;2(2):187-202.
54. Brasseur R, Divita G. Happy birthday cell penetrating peptides: already 20 years. Biochimica et biophysica acta. 2010;1798(12):2177-81. Epub 2010/09/10.
55. Richard JP, Melikov K, Vives E, Ramos C, Verbeure B, Gait MJ, et al. Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake. J Biol Chem. 2003;278(1):585-90. Epub 2002/11/02.
56. Trabulo S, Cardoso AS, Mano M, Pedroso De Lima MC. Cell-Penetrating Peptides—Mechanisms of Cellular Uptake and Generation of Delivery Systems. Pharmaceuticals 2010;3(4):961-93.
57. Mueller J, Kretzschmar I, Volkmer R, Boisguerin P. Comparison of Cellular Uptake Using 22 CPPs in 4 Different Cell Lines. Bioconjugate Chem. 2008;19(12):2363-74.
58. Tünnemann G, Cardoso MC. Cell-Penetrating Peptides— Uptake, Toxicity, and Applications. In: Castanho M, editor. Membrane-Active Peptides: Methods adn Results on Structure and Function: International University Line; 2010. p. 329-62.
59. Maiolo JR, Ferrer M, Ottinger EA. Effects of cargo molecules on the cellular uptake of arginine-rich cell-penetrating peptides. Biochimica et biophysica acta. 2005;1712(2):161-72. Epub 2005/06/07.
60. Henriques ST, Costa J, Castanho MA. Re-evaluating the role of strongly charged sequences in amphipathic cell-penetrating peptides: a fluorescence study using Pep-1. FEBS letters. 2005;579(20):4498-502. Epub 2005/08/09.
61. Ziegler A, Seelig J. Contributions of glycosaminoglycan binding and clustering to the biological uptake of the nonamphipathic cell-penetrating peptide WR9. Biochemistry. 2011;50(21):4650-64. Epub 2011/04/16.
62. Subrizi A, Tuominen E, Bunker A, Róg T, Antopolsky M, Urtti A. Tat(48-60) peptide amino acid sequence is not unique in its cell penetrating properties and cell-surface glycosaminoglycans inhibit its cellular uptake. Journal of Controlled Release. 2012;158(2):277-85.
63. Verdurmen WP, Wallbrecher R, Schmidt S, Eilander J, Bovee-Geurts P, Fanghanel S, et al. Cell surface clustering of heparan sulfate proteoglycans by amphipathic cell-penetrating peptides does not contribute to uptake. Journal of controlled release : official journal of the Controlled Release Society. 2013;170(1):83-91. Epub 2013/05/15.
64. Jarver P, Mager I, Langel U. In vivo biodistribution and efficacy of peptide mediated delivery. Trends Pharmacol Sci. 2010;31(11):528-35.
65. Melo MN, Ferre R, Castanho MA. Antimicrobial peptides: linking partition, activity and high membrane-bound concentrations. Nature reviews Microbiology. 2009;7(3):245-50. Epub 2009/02/17.
66. Deshayes S, Morris MC, Divita G, Heitz F. Interactions of amphipathic CPPs with model membranes. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2006;1758(3):328-35.
67. Teixeira V, Feio MJ, Bastos M. Role of lipids in the interaction of antimicrobial peptides with membranes. Prog Lipid Res. 2012;51(2):149-77.
68. Thoren PE, Persson D, Esbjorner EK, Goksor M, Lincoln P, Norden B. Membrane binding and translocation of cell-penetrating peptides. Biochemistry. 2004;43(12):3471-89. Epub 2004/03/24.
69. Bolhassani A. Potential efficacy of cell-penetrating peptides for nucleic acid and drug delivery in cancer. Bba-Rev Cancer. 2011;1816(2):232-46.
70. Antoniou X, Borsello T. Cell Permeable Peptides: A Promising Tool to Deliver Neuroprotective Agents in the Brain. Pharmaceuticals. 2010;3(2).

71. Edelstein ML, Abedi MR, Wixon J. Gene therapy clinical trials worldwide to 2007--an update. The journal of gene medicine. 2007;9(10):833-42. Epub 2007/08/28.
72. Scholz C, Wagner E. Therapeutic plasmid DNA versus siRNA delivery: common and different tasks for synthetic carriers. Journal of controlled release : official journal of the Controlled Release Society. 2012;161(2):554-65. Epub 2011/11/30.
73. Videira M, Arranja A, Rafael D, Gaspar R. Preclinical development of siRNA therapeutics: Towards the match between fundamental science and engineered systems. Nanomedicine : nanotechnology, biology, and medicine. 2013. Epub 2013/12/18.
74. Thierry AR, Abes S, Resina S, Travo A, Richard JP, Prevot P, et al. Comparison of basic peptides- and lipid-based strategies for the delivery of splice correcting oligonucleotides. Biochimica et biophysica acta. 2006;1758(3):364-74. Epub 2005/12/20.
75. Thompson JD. Clinical development of synthetic siRNA therapeutics. Drug Discovery Today: Therapeutic Strategies. 2013(0).
76. Rittner K, Benavente A, Bompard-Sorlet A, Heitz F, Divita G, Brasseur R, et al. New basic membrane-destabilizing peptides for plasmid-based gene delivery in vitro and in vivo. Molecular therapy : the journal of the American Society of Gene Therapy. 2002;5(2):104-14. Epub 2002/02/07.
77. Crombez L, Charnet A, Morris MC, Aldrian-Herrada G, Heitz F, Divita G. A non-covalent peptide-based strategy for siRNA delivery. Biochemical Society transactions. 2007;35(Pt 1):44-6. Epub 2007/01/20.
78. Meade BR, Dowdy SF. Exogenous siRNA delivery using peptide transduction domains/cell penetrating peptides. Advanced drug delivery reviews. 2007;59(2–3):134-40.
79. Kumar P, Wu H, McBride JL, Jung KE, Kim MH, Davidson BL, et al. Transvascular delivery of small interfering RNA to the central nervous system. Nature. 2007;448(7149):39-43. Epub 2007/06/19.
80. Moulton HM, Moulton JD. Morpholinos and their peptide conjugates: Therapeutic promise and challenge for Duchenne muscular dystrophy. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2010;1798(12):2296-303.
81. Sugita T, Yoshikawa T, Mukai Y, Yamanada N, Imai S, Nagano K, et al. Comparative study on transduction and toxicity of protein transduction domains. British journal of pharmacology. 2008;153(6):1143-52. Epub 2008/01/29.
82. Basheer M, Schwalb H, Shefler I, Levdansky L, Mekori YA, Gorodetsky R. Blood pressure modulation following activation of mast cells by cationic cell penetrating peptides. Peptides. 2011;32(12):2444-51.
83. Khafagy el S, Kamei N, Nielsen EJ, Nishio R, Takeda-Morishita M. One-month subchronic toxicity study of cell-penetrating peptides for insulin nasal delivery in rats. Eur J Pharm Biopharm. 2013;85(3 Pt A):736-43. Epub 2013/09/26.
84. Carter E, Lau CY, Tosh D, Ward SG, Mrsny RJ. Cell penetrating peptides fail to induce an innate immune response in epithelial cells in vitro: implications for continued therapeutic use. Eur J Pharm Biopharm. 2013;85(1):12-9. Epub 2013/08/21.
85. Saar K, Lindgren M, Hansen M, Eiriksdottir E, Jiang Y, Rosenthal-Aizman K, et al. Cell-penetrating peptides: A comparative membrane toxicity study. Anal Biochem. 2005;345(1):55-65.
86. El-Andaloussi S, Jarver P, Johansson HJ, Langel U. Cargo-dependent cytotoxicity and delivery efficacy of cell-penetrating peptides: a comparative study. Biochem J. 2007;407:285-92.
87. Desai PR, Cormier AR, Shah PP, Patlolla RR, Paravastu AK, Singh M. P solid-state NMR based monitoring of permeation of cell penetrating peptides into skin. Eur J Pharm Biopharm. 2013. Epub 2013/05/25.

88. Johnson LN, Cashman SM, Read SP, Kumar-Singh R. Cell penetrating peptide POD mediates delivery of recombinant proteins to retina, cornea and skin. Vision research. 2010;50(7):686-97. Epub 2009/09/08.
89. Patel EN, Wang J, Kim KJ, Borok Z, Crandall ED, Shen WC. Conjugation with Cationic Cell-Penetrating Peptide Increases Pulmonary Absorption of Insulin. Mol Pharm. 2009;6(2):492-503.
90. Cai SR, Xu G, Becker-Hapak M, Ma M, Dowdy SF, McLeod HL. The kinetics and tissue distribution of protein transduction in mice. Eur J Pharm Sci. 2006;27(4):311-9. Epub 2005/12/27.
91. Sarko D, Beijer B, Garcia Boy R, Nothelfer EM, Leotta K, Eisenhut M, et al. The pharmacokinetics of cell-penetrating peptides. Mol Pharm. 2010;7(6):2224-31. Epub 2010/09/18.
92. Phelan A, Elliott G, O'Hare P. Intercellular delivery of functional p53 by the herpesvirus protein VP22. Nat Biotechnol. 1998;16(5):440-3. Epub 1998/05/21.
93. Falanga A, Vitiello MT, Cantisani M, Tarallo R, Guarnieri D, Mignogna E, et al. A peptide derived from herpes simplex virus type 1 glycoprotein H: membrane translocation and applications to the delivery of quantum dots. Nanomedicine : nanotechnology, biology, and medicine. 2011;7(6):925-34. Epub 2011/06/15.
94. Langedijk JPM, Olijhoek T, Meloen RH. Application, efficiency and cargo-dependence of transport peptides. International Congress Series. 2005;1277(0):95-107.
95. Langedijk JP. Translocation activity of C-terminal domain of pestivirus Erns and ribotoxin L3 loop. J Biol Chem. 2002;277(7):5308-14. Epub 2001/10/24.
96. Oess S, Hildt E. Novel cell permeable motif derived from the PreS2-domain of hepatitis-B virus surface antigens. Gene therapy. 2000;7(9):750-8. Epub 2000/05/24.
97. Montrose K, Yang Y, Sun X, Wiles S, Krissansen GW. Xentry, a new class of cell-penetrating peptide uniquely equipped for delivery of drugs. Scientific reports. 2013;3:1661. Epub 2013/04/17.
98. Nakase I, Hirose H, Tanaka G, Tadokoro A, Kobayashi S, Takeuchi T, et al. Cell-surface accumulation of flock house virus-derived peptide leads to efficient internalization via macropinocytosis. Molecular therapy : the journal of the American Society of Gene Therapy. 2009;17(11):1868-76. Epub 2009/08/27.
99. Qi X, Droste T, Kao CC. Cell-penetrating peptides derived from viral capsid proteins. Molecular plant-microbe interactions : MPMI. 2011;24(1):25-36. Epub 2010/12/09.
100. Mudhakir D, Harashima H. Learning from the viral journey: how to enter cells and how to overcome intracellular barriers to reach the nucleus. The AAPS journal. 2009;11(1):65-77. Epub 2009/02/06.
101. Cheng S, Brooks CL, 3rd. Viral capsid proteins are segregated in structural fold space. PLoS computational biology. 2013;9(2):e1002905. Epub 2013/02/15.
102. Mbonye U, Karn J. Transcriptional control of HIV latency: Cellular signaling pathways, epigenetics, happenstance and the hope for a cure. Virology. 2014. Epub 2014/02/26.
103. Perera R, Kuhn RJ. Structural proteomics of dengue virus. Curr Opin Microbiol. 2008;11(4):369-77. Epub 2008/07/23.
104. Shukla D. Viral entry mechanisms: simplicity drives complexity. Febs J. 2009;276(24):7205. Epub 2009/11/03.
105. Hulo C, de Castro E, Masson P, Bougueleret L, Bairoch A, Xenarios I, et al. ViralZone: a knowledge resource to understand virus diversity. Nucleic Acids Res. 2011;39(Database issue):D576-82. Epub 2010/10/16.
106. Smith EC, Popa A, Chang A, Masante C, Dutch RE. Viral entry mechanisms: the increasing diversity of paramyxovirus entry. Febs J. 2009;276(24):7217-27. Epub 2009/11/03.
107. Marsh M, Helenius A. Virus entry: open sesame. Cell. 2006;124(4):729-40. Epub 2006/02/25.

108. Smith AE, Helenius A. How viruses enter animal cells. Science. 2004;304(5668):237-42. Epub 2004/04/10.
109. Last NB, Schlamadinger DE, Miranker AD. A common landscape for membrane-active peptides. Protein science : a publication of the Protein Society. 2013;22(7):870-82. Epub 2013/05/08.
110. Lins L, Decaffmeyer M, Thomas A, Brasseur R. Relationships between the orientation and the structural properties of peptides and their membrane interactions. Bba-Biomembranes. 2008;1778(7-8):1537-44.
111. Freire JM, Veiga AS, de la Torre BG, Santos NC, Andreu D, Da Poian AT, et al. Peptides as models for the structure and function of viral capsid proteins: Insights on dengue virus capsid. Peptide Science. 2013;100(4):325-36.
112. Mulder KC, Lima LA, Miranda VJ, Dias SC, Franco OL. Current scenario of peptide-based drugs: the key roles of cationic antitumor and antiviral peptides. Frontiers in microbiology. 2013;4:321. Epub 2013/11/08.
113. Gaspar D, Veiga AS, Castanho MA. From antimicrobial to anticancer peptides. A review. Frontiers in microbiology. 2013;4:294. Epub 2013/10/09.
114. Wadhwani P, Epand RF, Heidenreich N, Burck J, Ulrich AS, Epand RM. Membrane-active peptides and the clustering of anionic lipids. Biophys J. 2012;103(2):265-74. Epub 2012/08/03.
115. Kuhn RJ, Zhang W, Rossmann MG, Pletnev SV, Corver J, Lenches E, et al. Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. Cell. 2002;108(5):717-25. Epub 2002/03/15.
116. Mukhopadhyay S, Kuhn RJ, Rossmann MG. A structural perspective of the Flavivirus life cycle. Nature Reviews Microbiology. 2005;3(1):13-22.
117. Freire JM, Veiga AS, de la Torre BG, Andreu D, Castanho MA. Quantifying molecular partition of cell-penetrating peptide-cargo supramolecular complexes into lipid membranes: optimizing peptide-based drug delivery systems. J Pept Sci. 2013;19(4):182-9. Epub 2013/01/17.
118. Freire JM, Veiga AS, Rego de Figueiredo I, de la Torre BG, Santos NC, Andreu D, et al. Nucleic acid delivery by cell penetrating peptides derived from dengue virus capsid protein: design and mechanism of action. Febs J. 2014;281(1):191-215. Epub 2013/11/30.
119. Alves CS, Melo MN, Franquelim HG, Ferre R, Planas M, Feliu L, et al. Escherichia coli cell surface perturbation and disruption induced by antimicrobial peptides BP100 and pepR. J Biol Chem. 2010;285(36):27536-44. Epub 2010/06/23.
120. Freire JM, Veiga AS, Conceicao TM, Kowalczyk W, Mohana-Borges R, Andreu D, et al. Intracellular nucleic Acid delivery by the supercharged dengue virus capsid protein. Plos One. 2013;8(12):e81450. Epub 2013/12/18.
121. McNaughton BR, Cronican JJ, Thompson DB, Liu DR. Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(15):6111-6.
122. Thompson DB, Cronican JJ, Liu DR. Engineering and Identifying Supercharged Proteins for Macromolecule Delivery into Mammalian Cells. Method Enzymol. 2012;503:293-319.
123. Cronican JJ, Thompson DB, Beier KT, McNaughton BR, Cepko CL, Liu DR. Potent Delivery of Functional Proteins into Mammalian Cells in Vitro and in Vivo Using a Supercharged Protein. Acs Chem Biol. 2010;5(8):747-52.
124. Thompson DB, Villasenor R, Dorr BM, Zerial M, Liu DR. Cellular Uptake Mechanisms and Endosomal Trafficking of Supercharged Proteins. Chem Biol. 2012;19(7):831-43.
125. Cronican JJ, Beier KT, Davis TN, Tseng JC, Li WD, Thompson DB, et al. A Class of Human Proteins that Deliver Functional Proteins into Mammalian Cells In Vitro and In Vivo. Chem Biol. 2011;18(7):833-8.

126. Gautam A, Chaudhary K, Kumar R, Sharma A, Kapoor P, Tyagi A, et al. In silico approaches for designing highly effective cell penetrating peptides. J Transl Med. 2013;11:74. Epub 2013/03/23.
127. Frickey T, Lupas A. CLANS: a Java application for visualizing protein families based on pairwise similarity. Bioinformatics. 2004;20(18):3702-4. Epub 2004/07/31.
128. Magrane M, Consortium U. UniProt Knowledgebase: a hub of integrated protein data. Database : the journal of biological databases and curation. 2011;2011:bar009. Epub 2011/03/31.
129. Divita G. Bioactive cell-penetrating peptides: kill two birds with one stone. Chem Biol. 2010;17(7):679-80. Epub 2010/07/28.
130. Hansen M, Kilk K, Langel U. Predicting cell-penetrating peptides. Advanced drug delivery reviews. 2008;60(4-5):572-9. Epub 2007/11/30.
131. Holton TA, Pollastri G, Shields DC, Mooney C. CPPpred: prediction of cell penetrating peptides. Bioinformatics. 2013;29(23):3094-6.
132. abcam. Introduction to Flow Cytometry - A guide to flow cytometry. How does it work? What is it used for? 2014 [cited 2014]; Available from: http://www.abcam.com/index.html?pageconfig=resource&rid=11446.
133. Biosciences B. Introduction to Flow Cytometry: a learning guide. 2000; Available from: http://www.stemcell.umn.edu/prod/groups/med/@pub/@med/documents/asset/med_80691.pdf.
134. technologies MPid. LIVE/DEAD®Viability/Cytotoxicity Kit *for mammalian cells* 2005 [2014]; Available from: http://tools.lifetechnologies.com/content/sfs/manuals/mp03224.pdf.
135. So PT, Dong CY. Fluorescence Spectrophotometry. eLS. 2002.
136. Narhi LO, Cynthia HL, Ramachander R, Svitel J, Jiang Y. Optical Spectroscopic Methods for the Analysis of Biological Macromolecules. In: Allewell NM, Narhi LO, Rayment I, editors. Molecular Biophysics for the Life Sciences Springer; 2013.
137. Matos PM, Franquelim HG, Castanho MA, Santos NC. Quantitative assessment of peptide-lipid interactions. Ubiquitous fluorescence methodologies. Biochimica et biophysica acta. 2010;1798(11):1999-2012. Epub 2010/07/22.
138. Clarke RJ. The dipole potential of phospholipid membranes and methods for its detection. Advances in colloid and interface science. 2001;89-90:263-81. Epub 2001/02/24.
139. O'Shea P. Intermolecular interactions with/within cell membranes and the trinity of membrane potentials: kinetics and imaging. Biochemical Society transactions. 2003;31(Pt 5):990-6. Epub 2003/09/25.
140. Lambacher A, Fromherz P. Orientation of hemicyanine dye in lipid membrane measured by fluorescence interferometry on a silicon chip. J Phys Chem B. 2001;105(2):343-6.
141. Matos PM, Goncalves S, Santos NC. Interaction of peptides with biomembranes assessed by potential-sensitive fluorescent probes. J Pept Sci. 2008;14(4):407-15. Epub 2008/01/15.
142. Kelly SM, Jess TJ, Price NC. How to study proteins by circular dichroism. Bba-Proteins Proteom. 2005;1751(2):119-39.
143. Greenfield NJ. Applications of circular dichroism in protein and peptide analysis. TrAC Trends in Analytical Chemistry. 1999;18(4):236-44.
144. Domingues MM, Santiago PS, Castanho MARB, Santos NC. What can light scattering spectroscopy do for membrane-active peptide studies. J Pept Sci. 2008;14(4):394-400.
145. Nicolson GL. The Fluid-Mosaic Model of Membrane Structure: Still relevant to understanding the structure, function and dynamics of biological membranes after more than 40years. Biochimica et biophysica acta. 2013. Epub 2013/11/06.
146. van Meer G, Voelker DR, Feigenson GW. Membrane lipids: where they are and how they behave. Nature reviews Molecular cell biology. 2008;9(2):112-24. Epub 2008/01/25.

147. Pomorski TG, Nylander T, Cardenas M. Model cell membranes: Discerning lipid and protein contributions in shaping the cell. Advances in colloid and interface science. 2013. Epub 2013/11/26.
148. Nuno CS, Miguel ARBC. Lipossomas: a bala mágica acertou? SBQ. 2002;25(6B).
149. Zhao H, Lappalainen P. A simple guide to biochemical approaches for analyzing protein-lipid interactions. Molecular biology of the cell. 2012;23(15):2823-30. Epub 2012/08/01.
150. Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT. Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res. 2013;41(Web Server issue):W349-57. Epub 2013/06/12.
151. Gautier R, Douguet D, Antonny B, Drin G. HELIQUEST: a web server to screen sequences with specific alpha-helical properties. Bioinformatics. 2008;24(18):2101-2. Epub 2008/07/30.
152. Kaur H, Garg A, Raghava GP. PEPstr: a de novo method for tertiary structure prediction of small bioactive peptides. Protein and peptide letters. 2007;14(7):626-31. Epub 2007/09/28.
153. Gofman Y, Haliloglu T, Ben-Tal N. Monte Carlo simulations of peptide-membrane interactions with the MCPep web server. Nucleic Acids Res. 2012;40(Web Server issue):W358-63. Epub 2012/06/15.
154. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nature methods. 2012;9(7):671-5. Epub 2012/08/30.
155. Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. Proteins. 2011. Epub 2011/11/19.
156. Freire JM, Veiga AS, de la Torre BG, Santos NC, Andreu D, Da Poian AT, et al. Peptides as Models for the Structure and Function of Viral Capsid Proteins: Insights on Dengue Virus Capsid. Peptide Science2013.
157. Bahnsen JS, Franzyk H, Sandberg-Schaal A, Nielsen HM. Antimicrobial and cell-penetrating properties of penetratin analogs: effect of sequence and secondary structure. Biochimica et biophysica acta. 2013;1828(2):223-32. Epub 2012/10/23.
158. Walrant A, Vogel A, Correia I, Lequin O, Olausson BE, Desbat B, et al. Membrane interactions of two arginine-rich peptides with different cell internalization capacities. Biochimica et biophysica acta. 2012;1818(7):1755-63. Epub 2012/03/10.
159. Palm-Apergi C, Lonn P, Dowdy SF. Do Cell-Penetrating Peptides Actually "Penetrate" Cellular Membranes? Molecular Therapy. 2012;20(4):695-7.
160. Henriques ST, Castanho MA, Pattenden LK, Aguilar MI. Fast membrane association is a crucial factor in the peptide pep-1 translocation mechanism: a kinetic study followed by surface plasmon resonance. Biopolymers. 2010;94(3):314-22. Epub 2010/01/06.
161. Munoz-Morris MA, Heitz F, Divita G, Morris MC. The peptide carrier Pep-1 forms biologically efficient nanoparticle complexes. Biochemical and biophysical research communications. 2007;355(4):877-82. Epub 2007/03/03.
162. Franquelim HG, Veiga AS, Weissmuller G, Santos NC, Castanho MA. Unravelling the molecular basis of the selectivity of the HIV-1 fusion inhibitor sifuvirtide towards phosphatidylcholine-rich rigid membranes. Biochimica et biophysica acta. 2010;1798(6):1234-43. Epub 2010/02/16.
163. Gaspar D, Veiga AS, Sinthuvanich C, Schneider JP, Castanho MA. Anticancer peptide SVS-1: efficacy precedes membrane neutralization. Biochemistry. 2012;51(32):6263-5. Epub 2012/07/31.
164. Cézard C, Silva-Pires V, Mullié C, Sonnet P. Antibacterial Peptides: A Review. In: Mendez-Vilas A, editor. Science against Microbial Pathogens: Communicating Current Research and Technological Advances: Formatex Research Center; 2011.

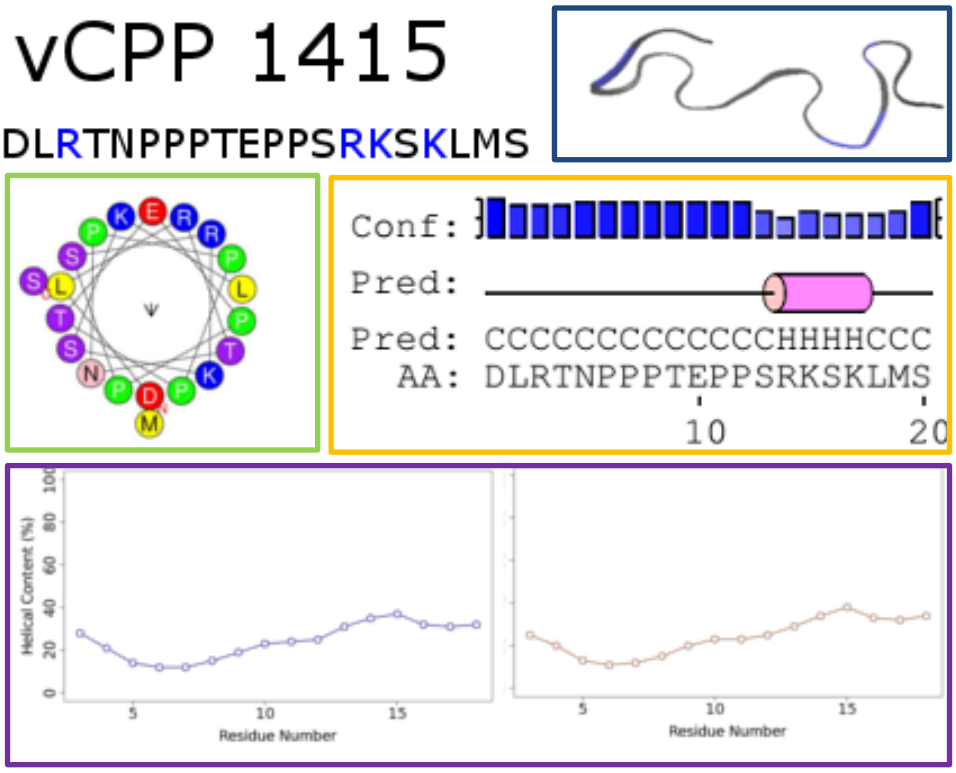
165. Li YM, Xiang Q, Zhang QH, Huang YD, Su ZJ. Overview on the recent study of antimicrobial peptides: Origins, functions, relative mechanisms and application. Peptides. 2012;37(2):207-15.
166. Melo MN, Castanho MA. The Mechanism of Action of Antimicrobial Peptides: Lipid Vesicles vs. Bacteria. Frontiers in immunology. 2012;3:236. Epub 2012/08/10.
167. Henriques ST, Melo MN, Castanho MA. Cell-penetrating peptides and antimicrobial peptides: how different are they? Biochem J. 2006;399(1):1-7. Epub 2006/09/08.
168. Nekhotiaeva N, Elmquist A, Rajarao GK, Hallbrink M, Langel U, Good L. Cell entry and antimicrobial properties of eukaryotic cell-penetrating peptides. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2004;18(2):394-6. Epub 2003/12/06.
169. Splith K, Neundorf I. Antimicrobial peptides with cell-penetrating peptide properties and vice versa. European biophysics journal : EBJ. 2011;40(4):387-97. Epub 2011/02/22.
170. Bobone S, Piazzon A, Orioni B, Pedersen JZ, Nan YH, Hahm KS, et al. The thin line between cell-penetrating and antimicrobial peptides: the case of Pep-1 and Pep-1-K. J Pept Sci. 2011;17(5):335-41.
171. Antunes E, Azoia NG, Matama T, Gomes AC, Cavaco-Paulo A. The activity of LE10 peptide on biological membranes using molecular dynamics, in vitro and in vivo studies. Colloids and surfaces B, Biointerfaces. 2013;106:240-7. Epub 2013/02/26.
172. Rodriguez-Rojas A, Rodriguez-Beltran J, Couce A, Blazquez J. Antibiotics and antibiotic resistance: A bitter fight against evolution. International journal of medical microbiology : IJMM. 2013. Epub 2013/03/23.
173. Theuretzbacher U. Global antibacterial resistance: The never-ending story. Journal of Global Antimicrobial Resistance. 2013;1(2):63-9.
174. Park SC, Park Y, Hahm KS. The Role of Antimicrobial Peptides in Preventing Multidrug-Resistant Bacterial Infections and Biofilm Formation. Int J Mol Sci. 2011;12(9):5971-92.
175. Lata S, Sharma BK, Raghava GP. Analysis and prediction of antibacterial peptides. BMC bioinformatics. 2007;8:263. Epub 2007/07/25.
176. Torrent M, Di Tommaso P, Pulido D, Nogues MV, Notredame C, Boix E, et al. AMPA: an automated web server for prediction of protein antimicrobial regions. Bioinformatics. 2012;28(1):130-1. Epub 2011/11/05.
177. Piotto SP, Sessa L, Concilio S, Iannelli P. YADAMP: yet another database of antimicrobial peptides. International journal of antimicrobial agents. 2012;39(4):346-51. Epub 2012/02/14.
178. Wang G, Li X, Wang Z. APD2: the updated antimicrobial peptide database and its application in peptide design. Nucleic Acids Res. 2009;37(Database issue):D933-7. Epub 2008/10/30.
179. Thomas S, Karnik S, Barai RS, Jayaraman VK, Idicula-Thomas S. CAMP: a useful resource for research on antimicrobial peptides. Nucleic Acids Res. 2010;38(Database issue):D774-80. Epub 2009/11/20.
180. Zhao X, Wu H, Lu H, Li G, Huang Q. LAMP: A Database Linking Antimicrobial Peptides. Plos One. 2013;8(6):e66557. Epub 2013/07/05.
181. Seshadri Sundararajan V, Gabere MN, Pretorius A, Adam S, Christoffels A, Lehvaslaiho M, et al. DAMPD: a manually curated antimicrobial peptide database. Nucleic Acids Res. 2012;40(Database issue):D1108-12. Epub 2011/11/24.
182. Torrent M, Di Tommaso P, Pulido D, Nogues MV, Notredame C, Boix E, et al. AMPA: an automated web server for prediction of protein antimicrobial regions. Bioinformatics. 2012;28(1):130-1.

183. Joseph S, Karnik S, Nilawe P, Jayaraman VK, Idicula-Thomas S. ClassAMP: a prediction tool for classification of antimicrobial peptides. IEEE/ACM transactions on computational biology and bioinformatics / IEEE, ACM. 2012;9(5):1535-8. Epub 2012/06/27.
184. Lata S, Mishra NK, Raghava GP. AntiBP2: improved version of antibacterial peptide prediction. BMC bioinformatics. 2010;11 Suppl 1:S19. Epub 2010/03/05.
185. Tyagi A, Kapoor P, Kumar R, Chaudhary K, Gautam A, Raghava GPS. In Silico Models for Designing and Discovering Novel Anticancer Peptides. Scientific reports. 2013;3.
186. Xiao X, Wang P, Lin WZ, Jia JH, Chou KC. iAMP-2L: A two-level multi-label classifier for identifying antimicrobial peptides and their functional types. Analytical Biochemistry. 2013;436(2):168-77.

Annexes

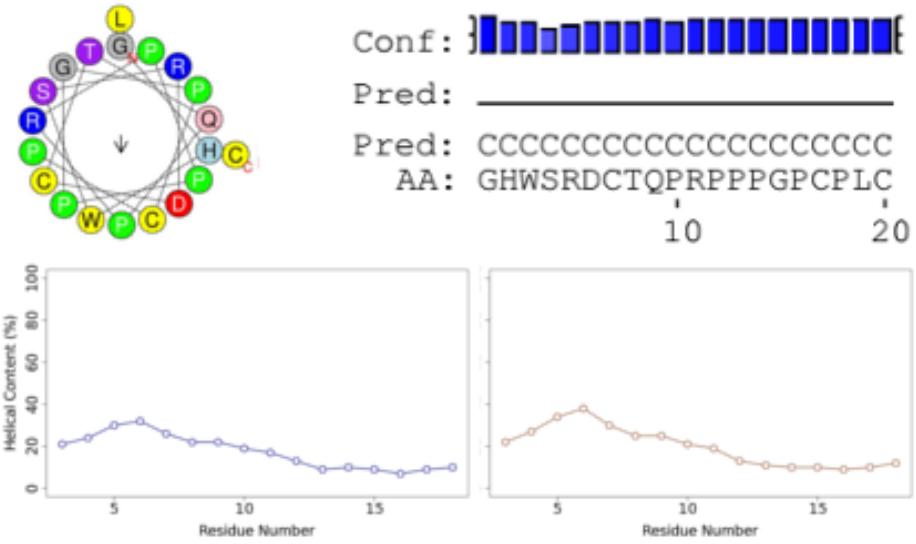
Annex A – viralCPP simulations

Computational analyses were conducted using several online tools. PSIPRED (150) was used to predict viralCPP propensity for secondary structure formation (*Conf* bars represent confidence of prediction per residue; Pink cylinders mark helical propensity, Yellow arrows mark β-sheet propensity, Black lines mark coiled regions). Heliquest (151) was used to construct helical projections of the sequences with a biochemical colour coding of the amino acid residues (Basic – Blue; Acidic – Red; Proline – Green; Highly Hydrophobic – Yellow; Hydrophobic– Grey; Histidine – Light Blue; Neutral – Purple) and indication of mean amphipathic moment vector. Illustrative tertiary structural models were computed by PEPstr (152). Monte Carlo simulations were run on the MCpep server (153) for prediction of viralCPP helical content in aqueous (blue line) and 20% anionic lipid membrane environments (brown line).



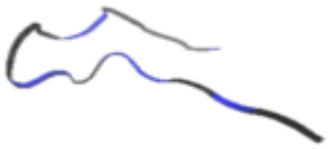
vCPP 0956

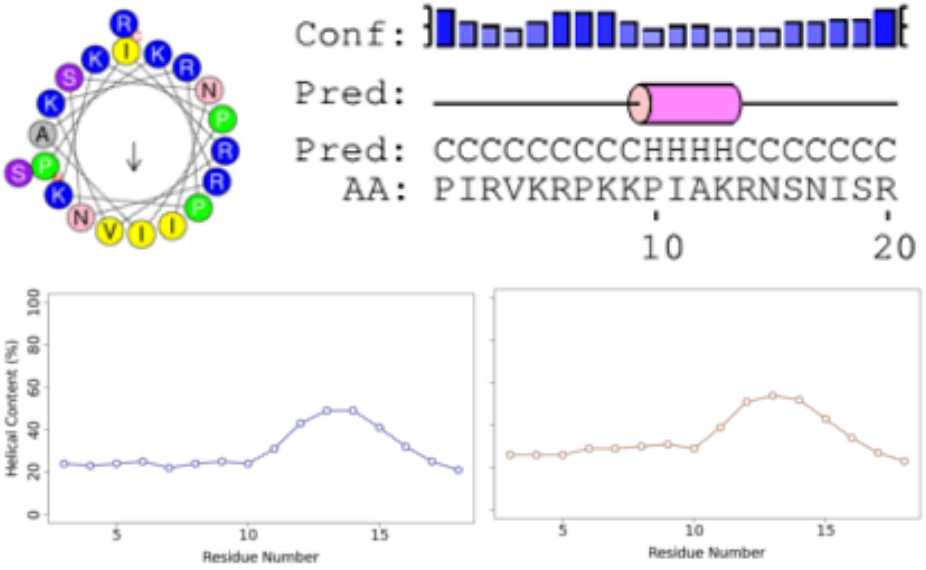
GHWSRDCTQPRPPPGPCPLC



vCPP 1396

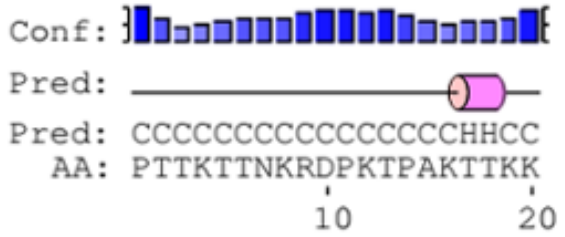
PIRVKRPKKPIAKRNSNISR

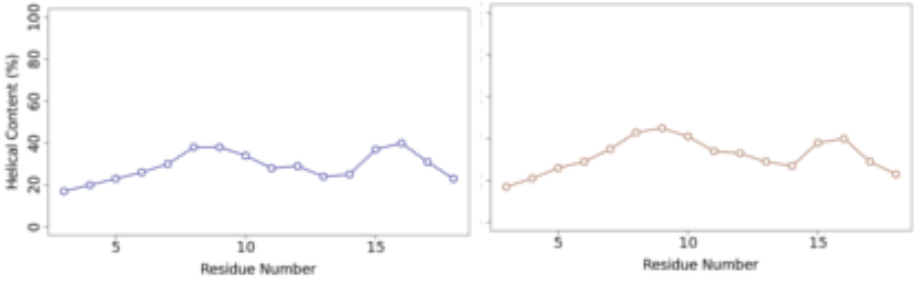




vCPP 1203

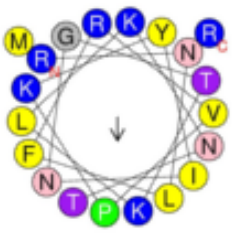
PTTKTTNKRDPKTPAKTTKK

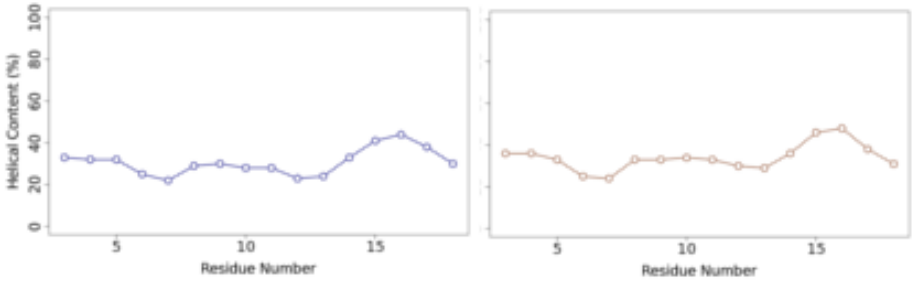




vCPP 0554

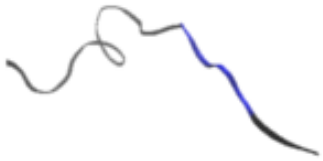
RNLFRVPKYINGTKLKNTMR

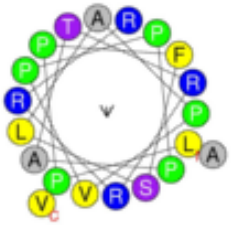


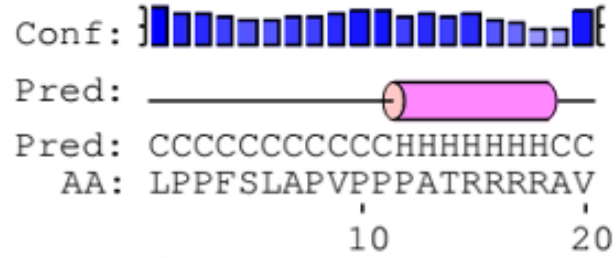


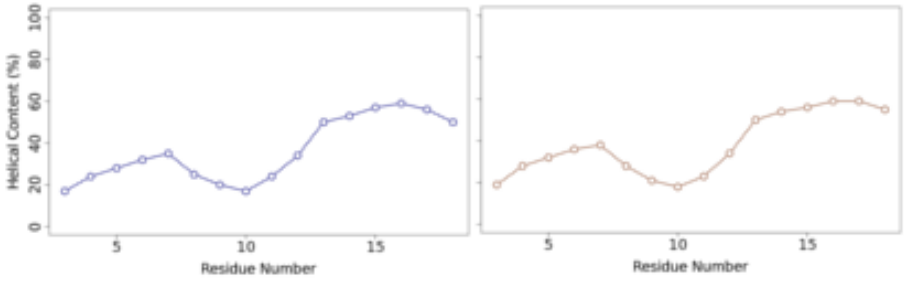
vCPP 0979

LPPFSLAPVPPPATRRRRAV



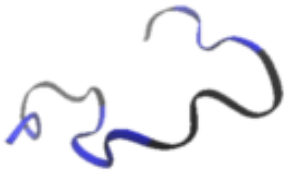


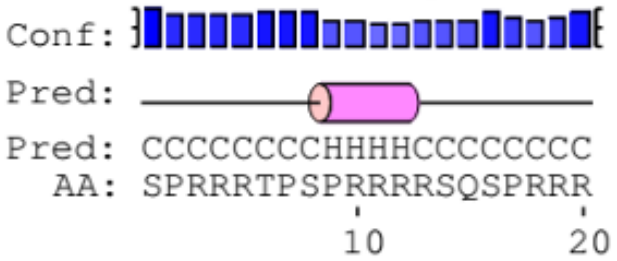


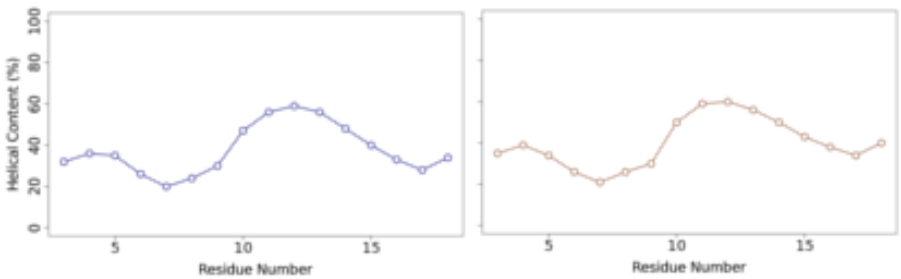


vCPP 0417

SPRRRTPSPRRRRSQSPRRR

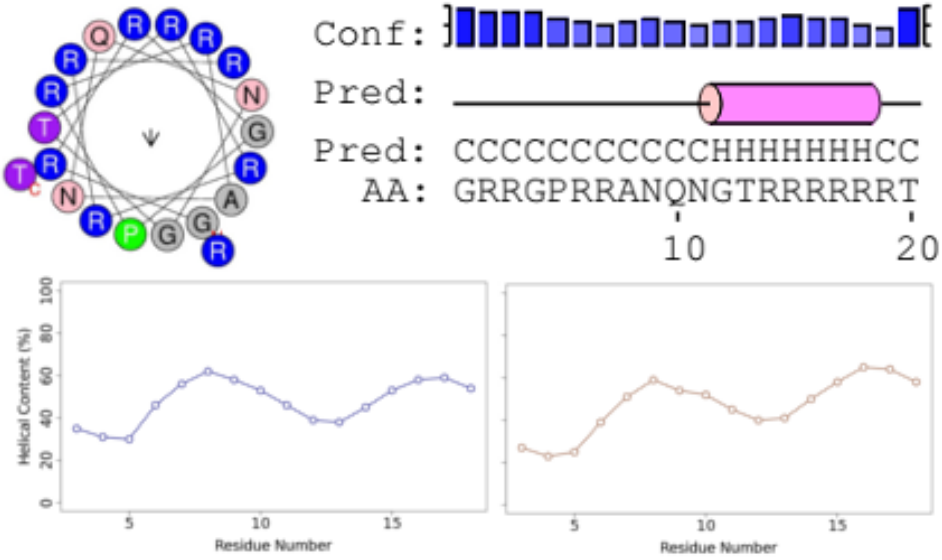






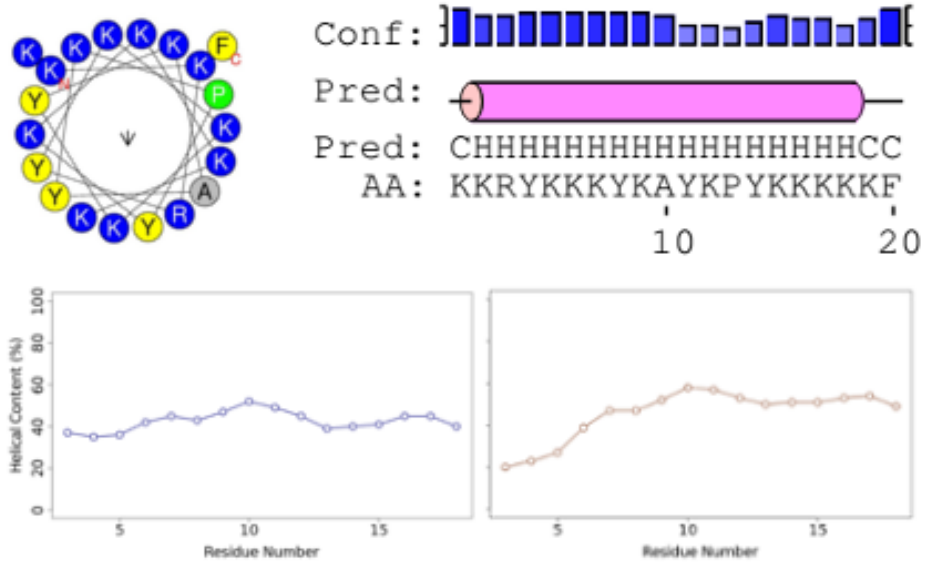
vCPP 1779

GRRGPRRANQNGTRRRRRRT



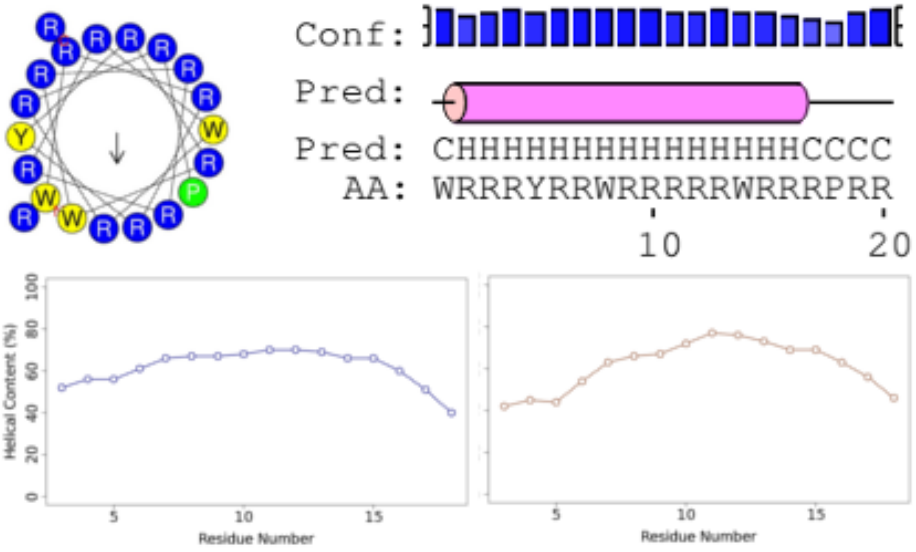
vCPP 0275

KKRYKKKYKAYKPYKKKKKF



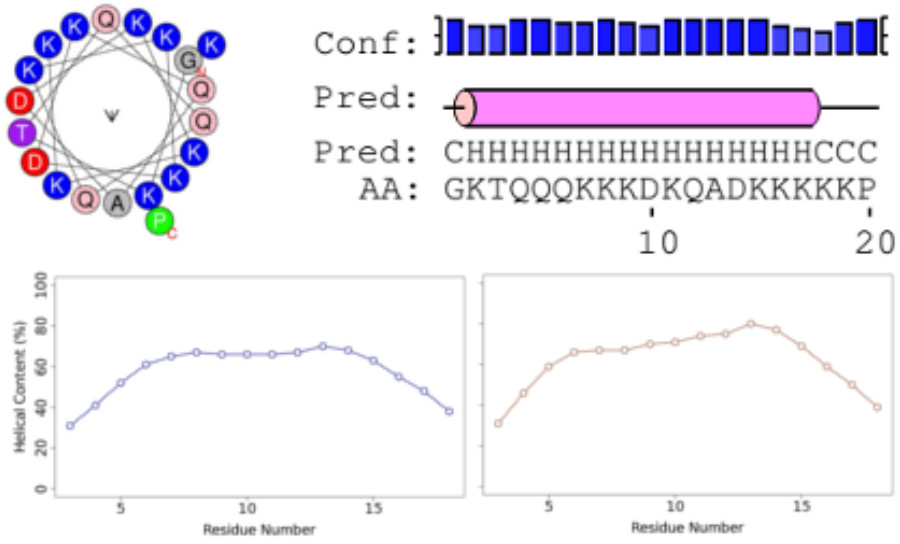
vCPP 2319

WRRRYRRWRRRRRWRRRPRR



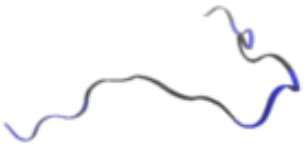
vCPP 1721

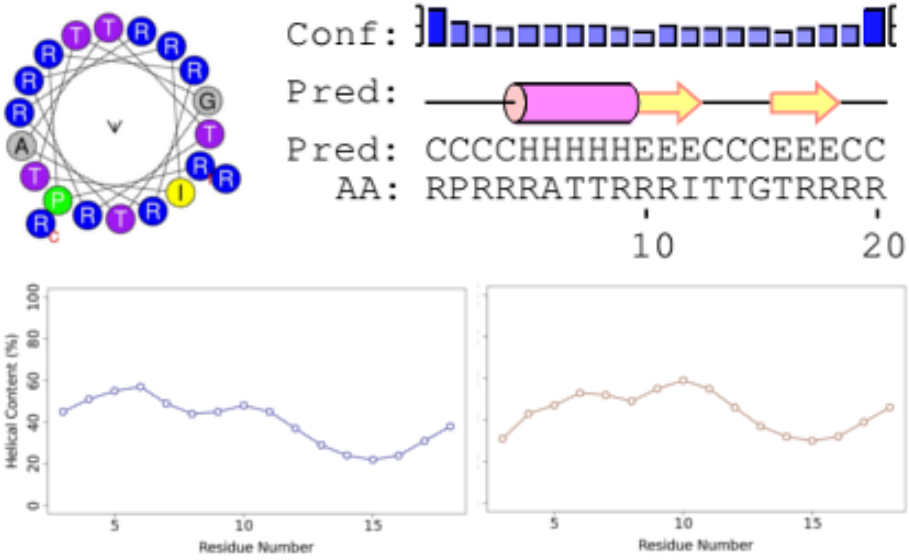
GKTQQQKKKDKQADKKKKKP



vCPP 0667

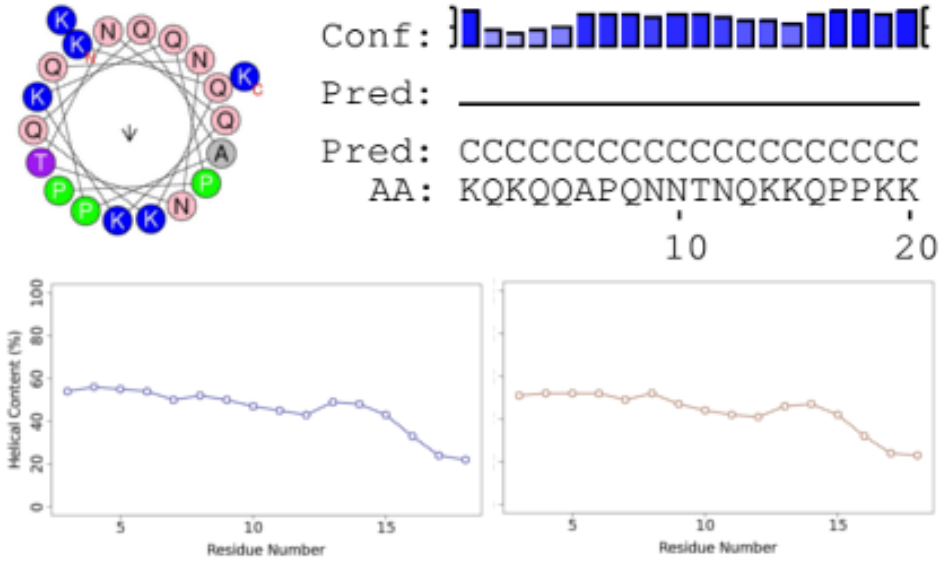
RPRRRATTRRRITTGTRRRR





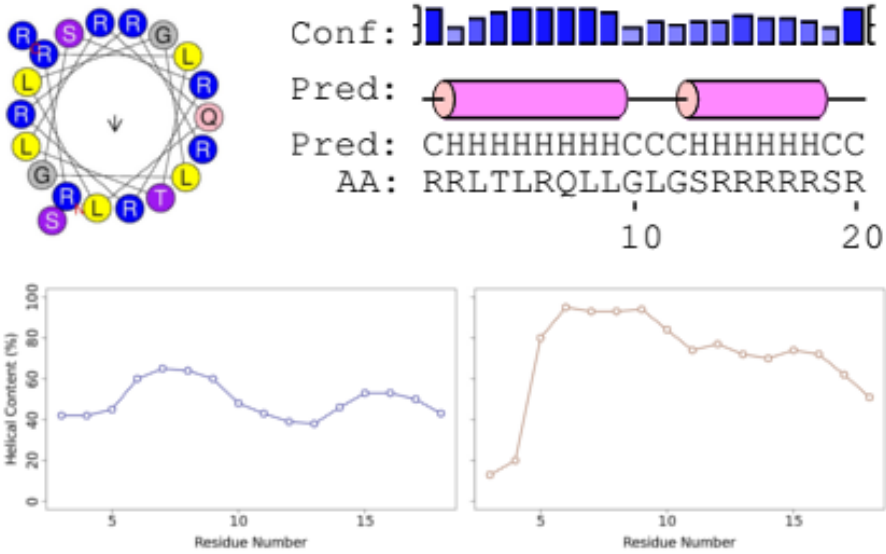
vCPP 1754

KQKQQAPQNNTNQKKQPPKK



vCPP 0769

RRLTLRQLLGLGSRRRRRSR



Annex B – Supplementary data

B.1 – Fluorescence spectroscopy

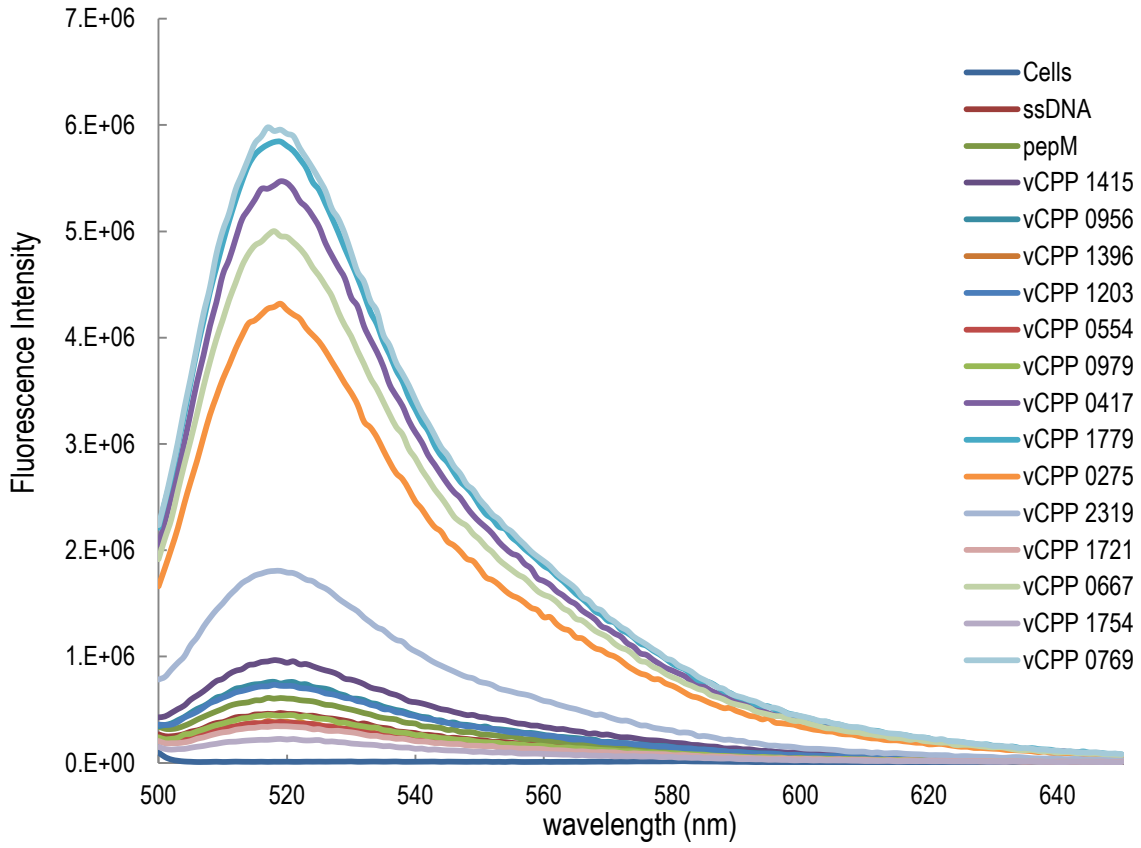


Figure B.1 – A-488 emission spectra collected for a replicate of the viralCPP/ssDNA delivery test. The underlying area of each curve was computed to quantify intensity of fluorescence.

B2 – Microplate reader and flow cytometry

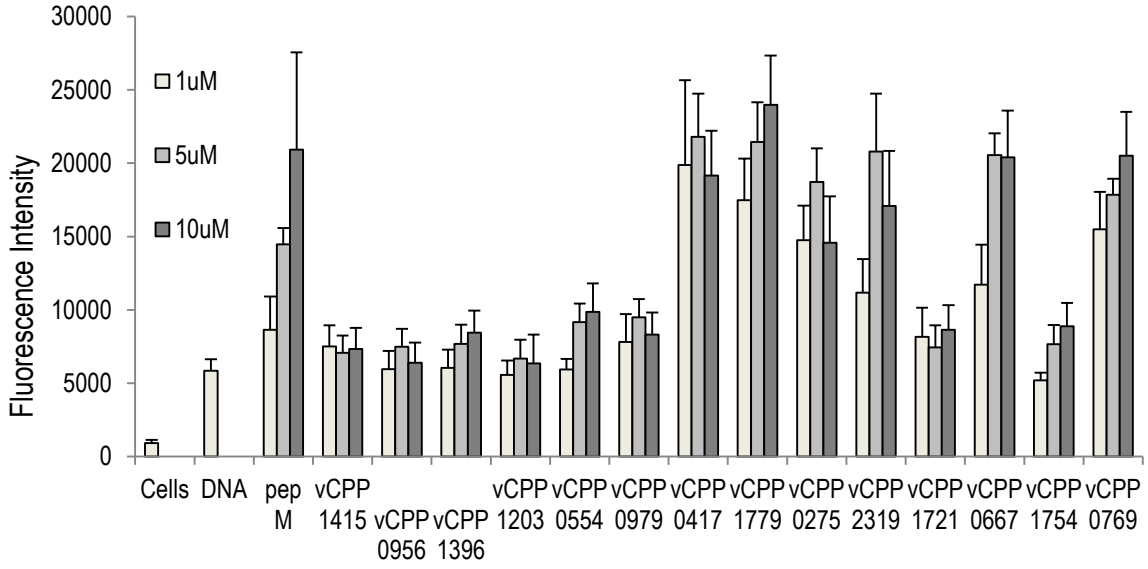


Figure B.2.1 – Mean of fluorescence measurements obtained with a PR (n = 5, SEM)

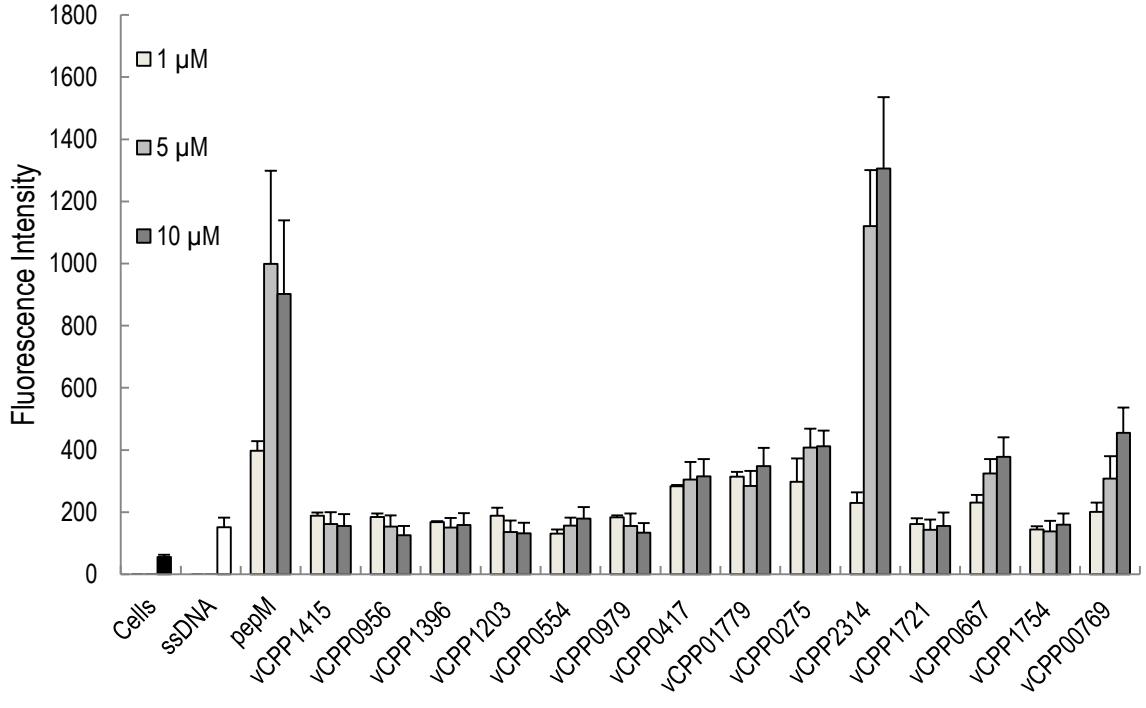


Figure B.2.2 – Mean of FC fluorescence intensity measurements (n=5, SEM) computed through the underlying area of (Intensity vs # events) histograms

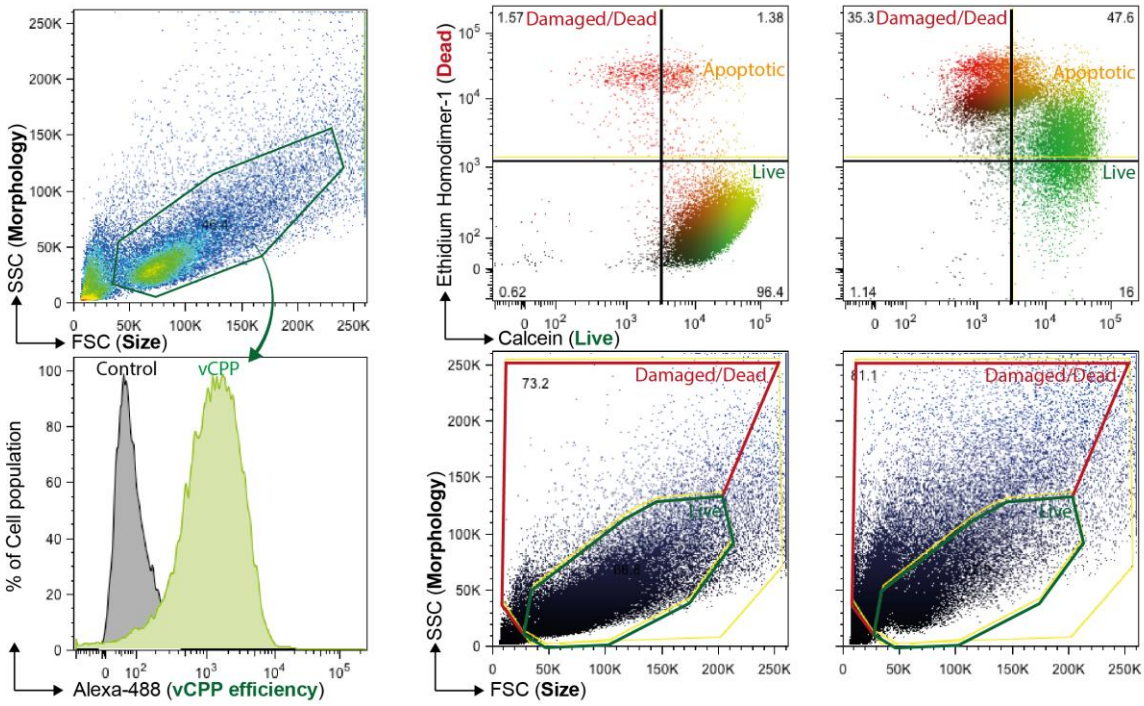


Figure B.2.3 – (Left panel) FC results from viralCPP mediated uptake of ssDNA-A488. Scatterplot of forward (FSC) and side-scattered (SSC) light displaying a gated population of cells in which delivery was successful. (Right panel) Live/dead assay of a sample of live cells, mostly stained with calcein, and of dead cells stained with ethidium. Apoptotic cells are stained by both dyes. Plots of FSC vs SSC can also be used to estimate cell death by gating cells displaying healthy morphology.

B3 – Confocal microscopy

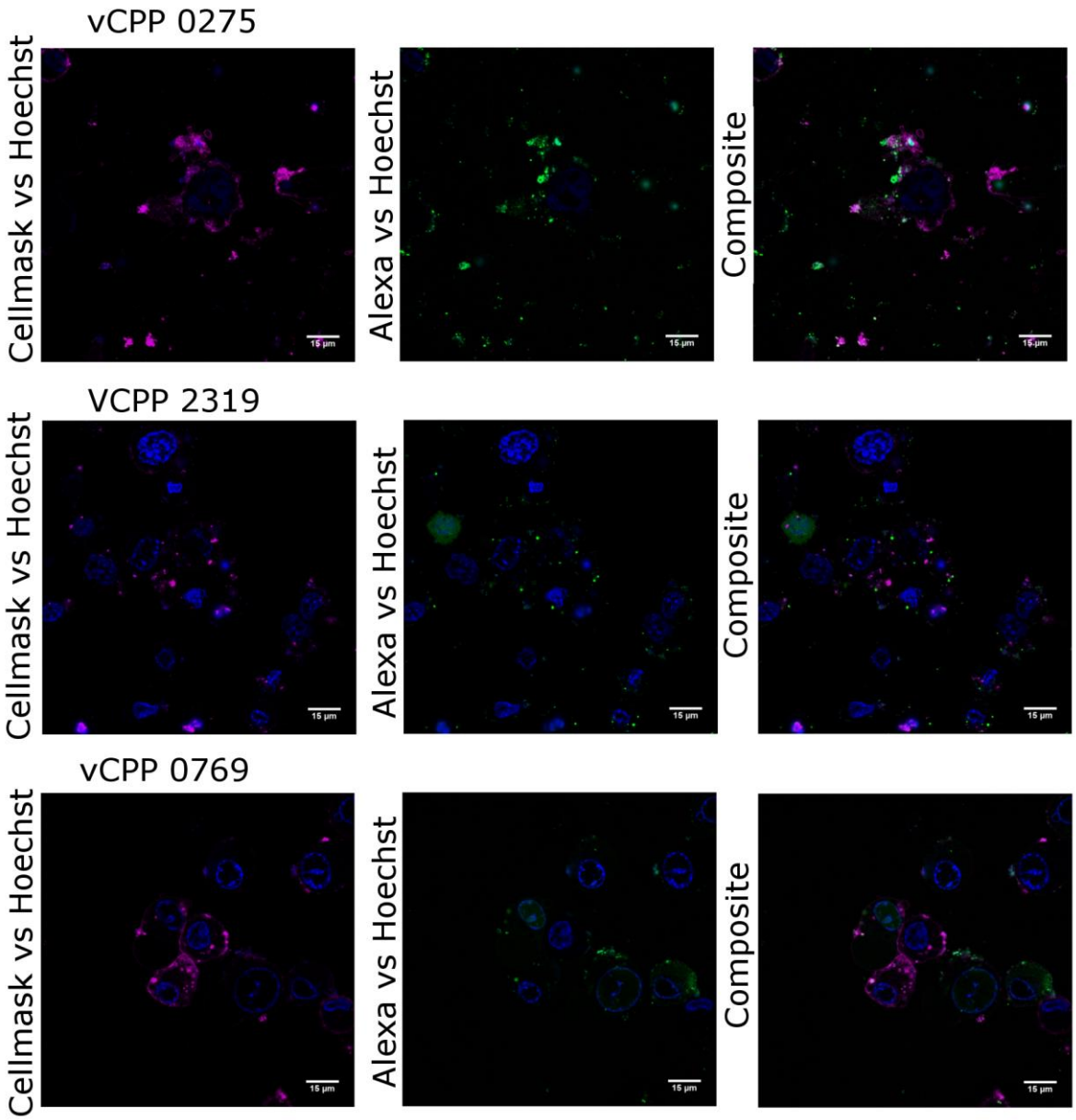


Figure B.3 – Examples of erratic Cellmask Deep Red membrane staining in CM images. Staining of membranes was successful in the panels related to viralCPP 0275, but was negligible with viralCPP 2319 samples. In the row further down (viralCPP 0769) only two of the cells in the observation field could be conveniently stained but did not coincide with peptide signal.

B4 – Membrane dipole potential sensing

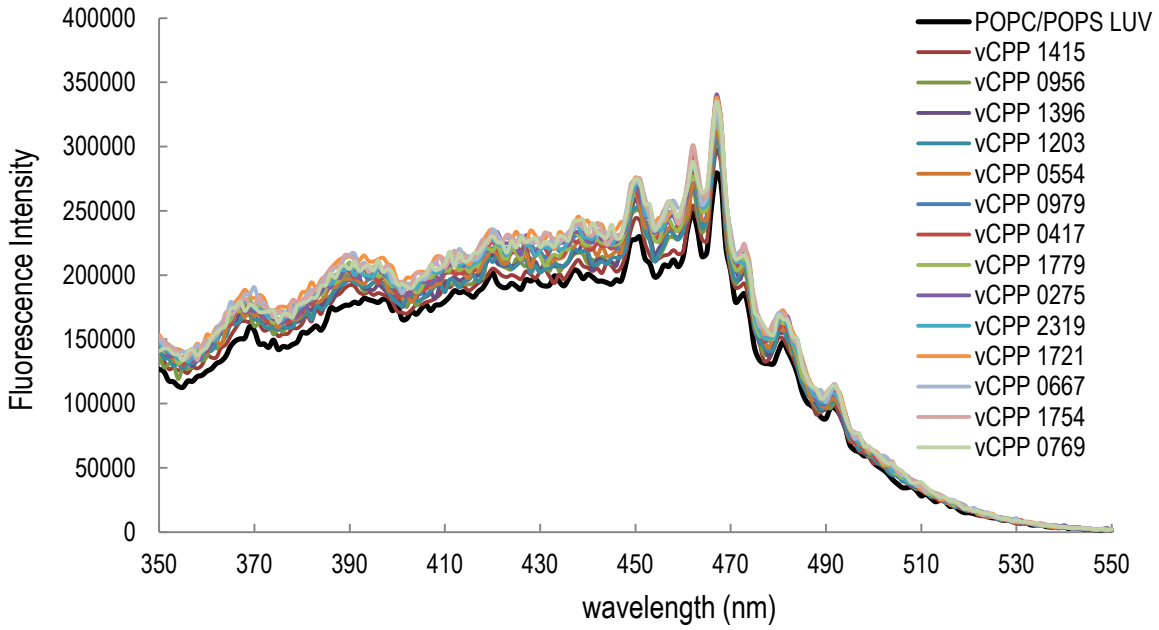


Figure B.4.1 – Excitation spectra of di-8-ANEPPS in LUVs before (black line) and after peptide additions. Upon normalization of each measurement, differential curves are obtained by subtracting the black spectrum to each of the colored spectra.

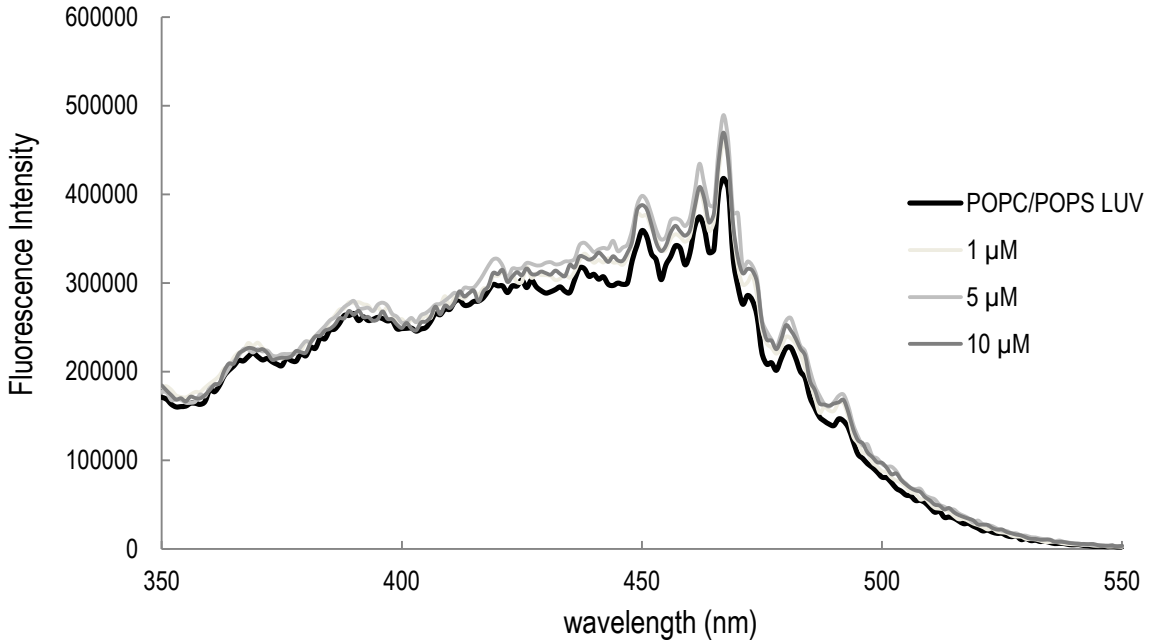


Figure B.4.2 - Excitation spectra of di-8-ANEPPS in LUVs before (black line) and after viralCPP 0769 additions at increasing concentration. Intensity values from these spectra are used to compute **R** and $\mathbf{R}_0$ upon determination of the wavelengths at which sinusoidal crests occur.

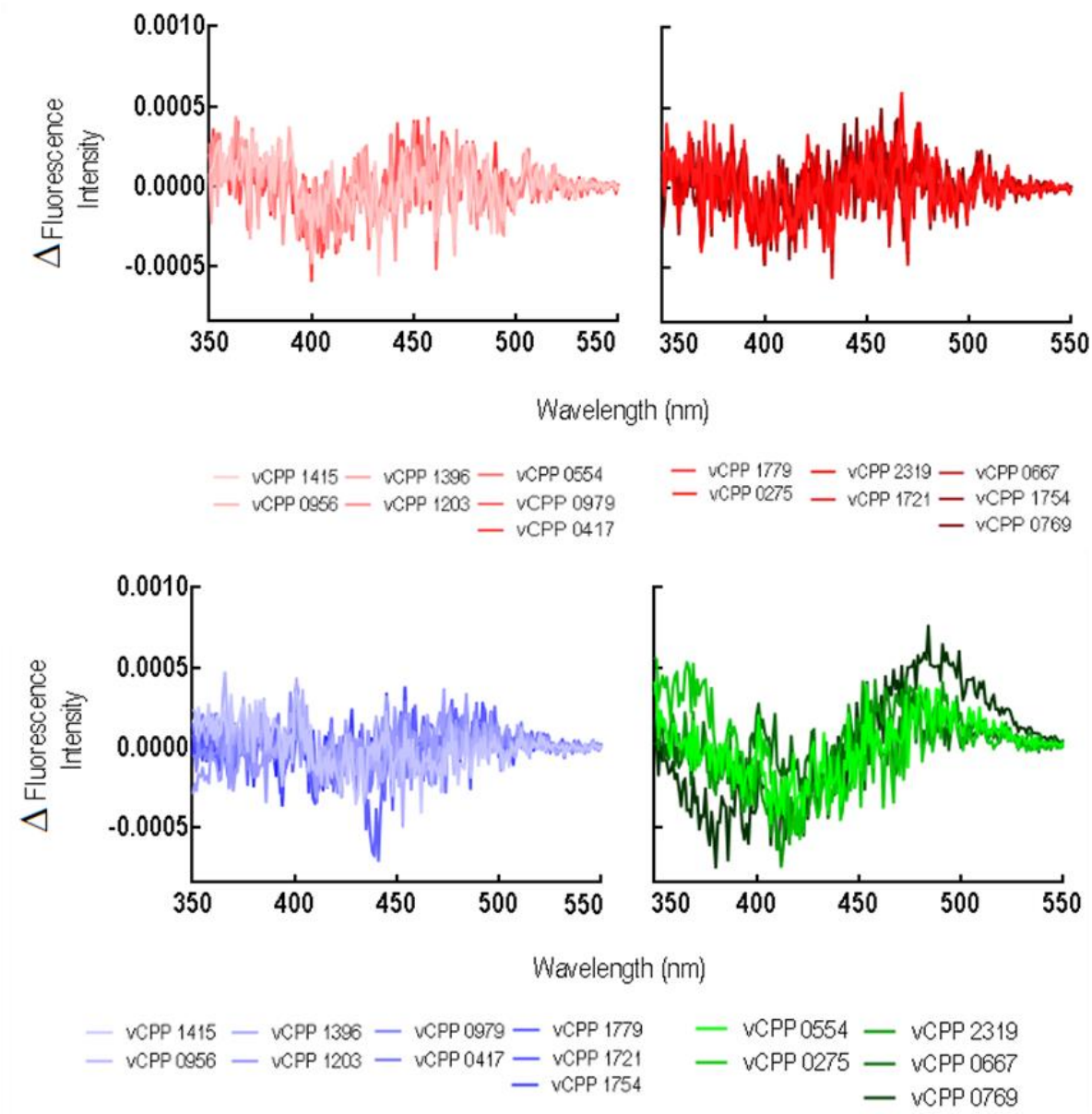


Figure B.4.3 - (Upper panels) Differential spectra of interactions reported by di-8-ANEPPS in POPC LUVs, viralCPP are split into two sets in shades of red. (Lower panels) Differential spectra of interactions reported by di-8-ANEPPS in POPC:POPS (4:1) LUVs, blue shaded set represents viralCPP without membrane affinity, while shades of green stand for high degree of viralCPP membrane interaction.

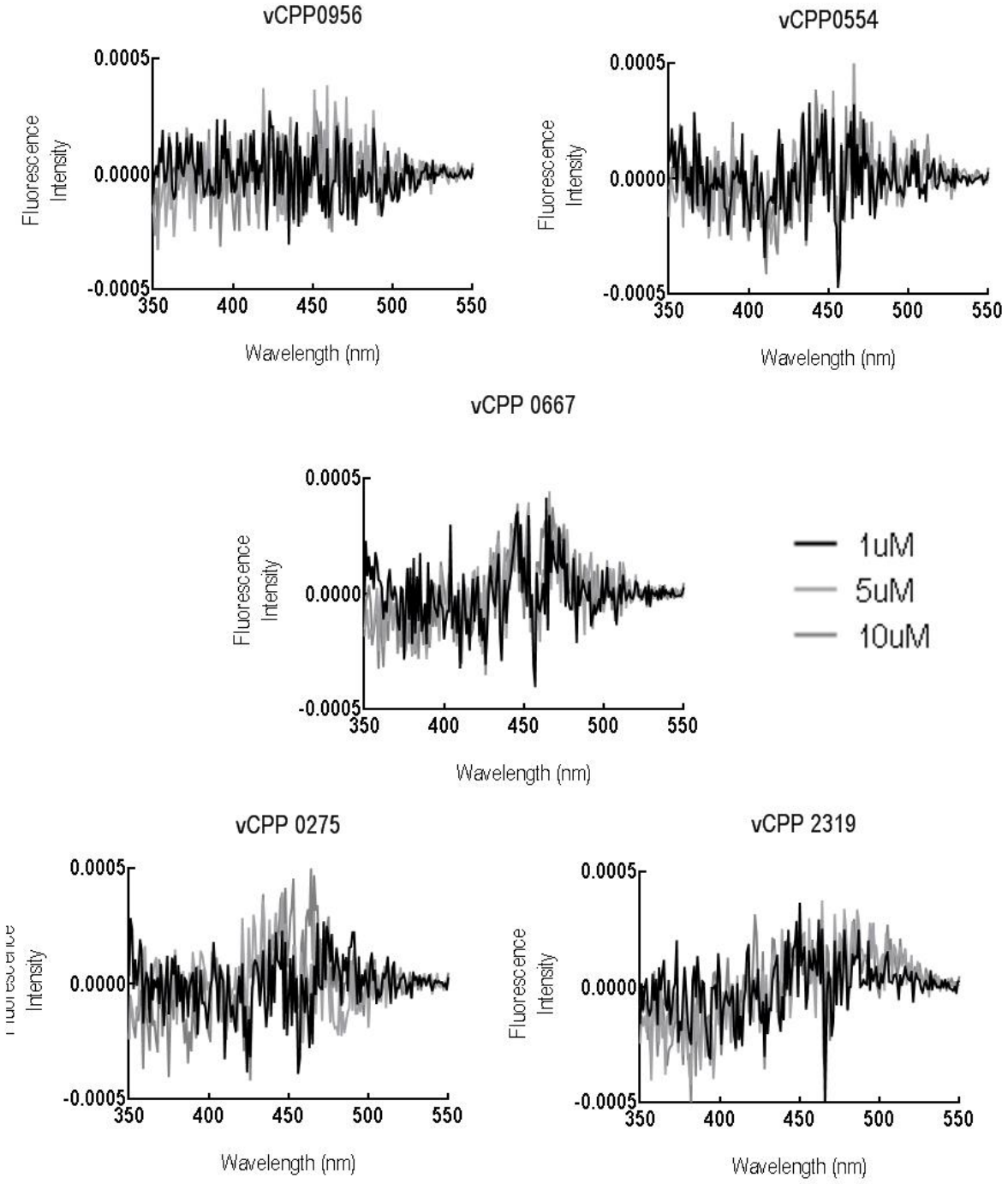


Figure B.4.4 - Illustration of shifts in amplitude of viralCPP differential spectra due to increases in peptide concentration. viralCPP 0956 represented as a negative control. viralCPP 0275 membrane perturbations could not be reported at concentrations bellow 10 µM. $\mathbf{R}/\mathbf{R}_0$ was computed for viralCPPs 0554, 0667 and 2319.

B5 – Dynamic light scattering

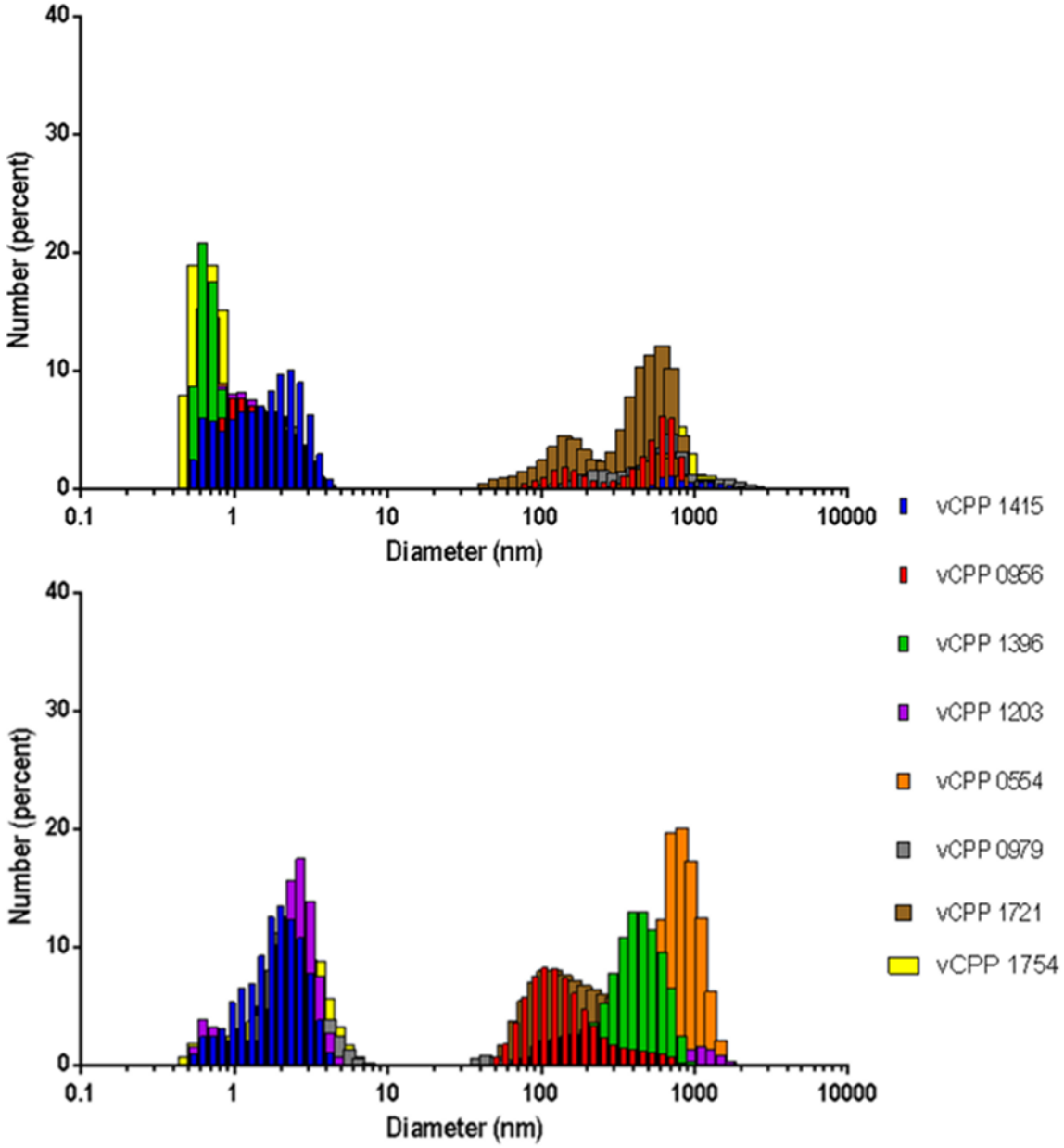


Figure B.5.1 – Hydrodynamic diameter of poor viralCPP leads before (upper panel) and after ssDNA conjugation (lower pannel) n=2

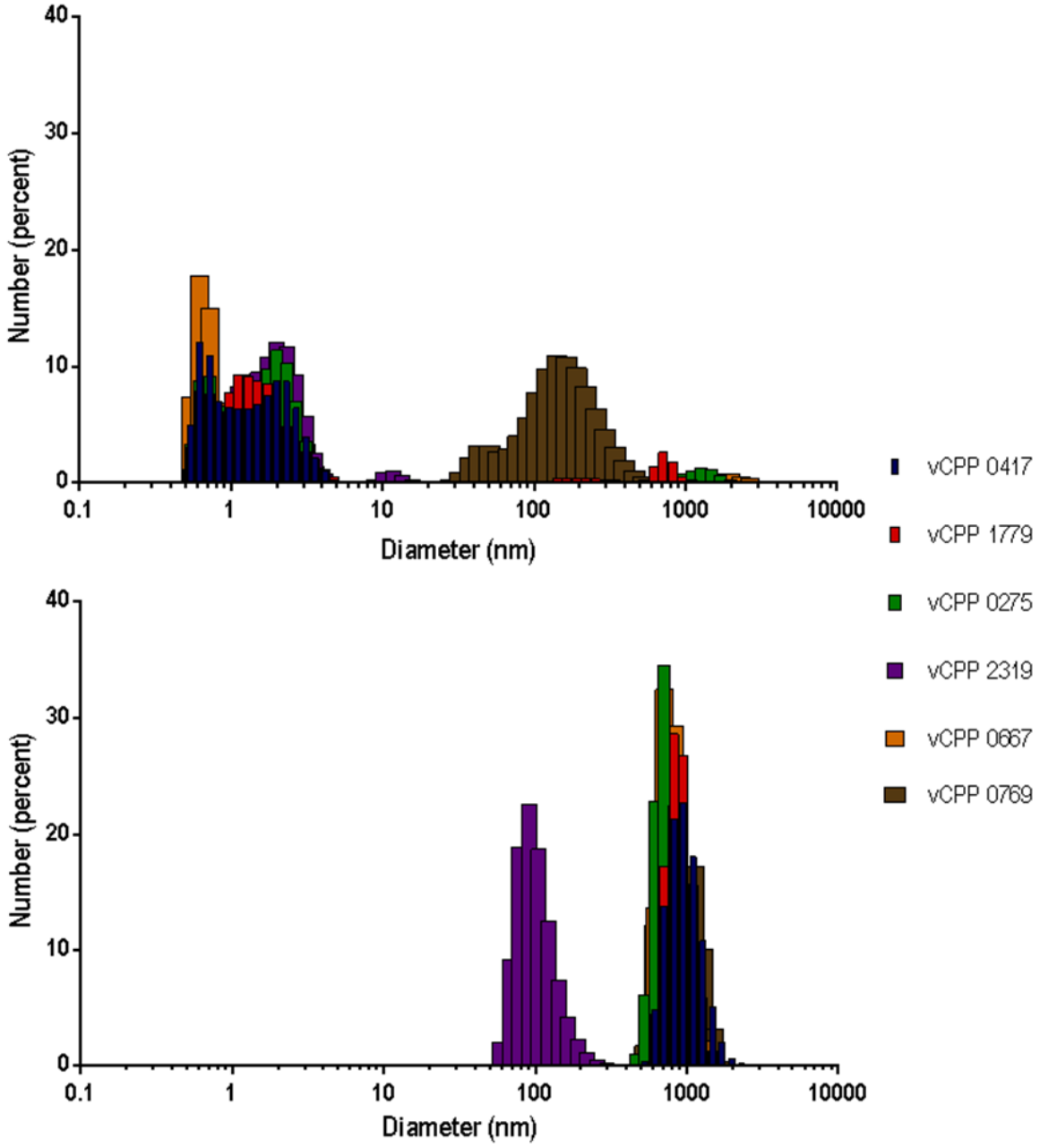


Figure B.5.2 – Hydrodynamic diameter of good viralCPP leads before (upper panel) and after ssDNA conjugation (lower pannel) n=2