



HAL
open science

Prediction of amphiphilic cell-penetrating peptide building blocks from protein-derived amino acid sequences for engineering of drug delivery nanoassemblies

Angelina Angelova, Borislav Angelov, Guillaume Feger

► **To cite this version:**

Angelina Angelova, Borislav Angelov, Guillaume Feger. Prediction of amphiphilic cell-penetrating peptide building blocks from protein-derived amino acid sequences for engineering of drug delivery nanoassemblies. *Journal of Physical Chemistry B*, 2020. hal-03093539

HAL Id: hal-03093539

<https://hal.science/hal-03093539>

Submitted on 4 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Prediction of Amphiphilic Cell-Penetrating Peptide
Building Blocks from Protein-Derived Amino Acid
Sequences for Engineering of Drug Delivery
Nanoassemblies

Guillaume Feger[§], Borislav Angelov[#] and Angelina Angelova^{§,}*

[§]Université Paris-Saclay, CNRS, Institut Galien Paris-Saclay UMR8612, F-92296 Châtenay-Malabry, France;

[#]Institute of Physics, ELI Beamlines, Academy of Sciences of the Czech Republic, Na Slovance 2, CZ-18221 Prague, Czech Republic.

ABSTRACT

Amphiphilic molecules, forming self-assembled nanoarchitectures, are typically composed of hydrophobic and hydrophilic domains. Peptide amphiphiles can be designed from two, three or four building blocks imparting novel structural and functional properties and affinities for interaction with cellular membranes or intracellular organelles. Here we present a combined numerical approach to design amphiphilic peptide scaffolds that are derived from the human nuclear Ki-67 protein. Ki-67 acts, like a biosurfactant, as a steric and electrostatic charge barrier against the collapse of mitotic chromosomes. The proposed predictive design of new Ki-67 protein-derived amphiphilic aminoacid sequences exploits the computational outcomes of a set of web-accessible predictors, which are based on machine learning methods. The ensemble of such artificial intelligence algorithms, involving support vector machine (SVM), random forest (RF) classifiers and neural networks (NN), enables the nano-engineering of a broad range of innovative peptide materials for therapeutic delivery in various applications. Amphiphilic cell-penetrating peptides (CPP), derived from natural protein sequences, may spontaneously form self-assembled nanocarriers characterized by enhanced cellular uptake. Thanks to their inherent low immunogenicity, they may enable the safe delivery of therapeutic molecules across the biological barriers.

Introduction

The development of novel drug delivery systems to improve the bioavailability of therapeutic compounds and reduce severe side effects is a research area of growing interest.¹⁻²⁴ Drug delivery nano-assemblies are composed of either lipids, peptides, surfactants, carbohydrates or polymers besides of different combinations of multiple amphiphilic ingredients.²⁵⁻³⁹ Screening of new carrier structures by application of machine learning in materials science⁴⁰⁻⁴⁹ creates alternatives to the complex synthesis of new drug molecules, which require a long and expensive process of biological assays and physico-chemical characterization of the formulations up to clinical trials.

The focus of the present work is on bio-inspired peptides with self-assembly properties as designed candidates for more efficient and safe drug delivery nanocarriers (Figure 1). Cell penetrating peptides (CPP) are short sequences of amino acids (generally up to 35 aa), which have a capacity to cross the cellular bilayer membranes without damaging them.⁵⁰⁻⁶⁰ The research on CPPs has been inspired by the fact that proteins (like transcription-activating factors, signal sequences and homeoproteins) involve subunits that may facilitate their translocation across the plasma membranes into the inner cellular compartments.^{50-52,61-63} This finding has stimulated a number of investigations on the design of synthetic aminoacid sequences, which self-assemble and enable the cellular internalization of various compounds.^{3-7,53-60,64-84} The large diversity of CPPs has been classified into sequences of cationic, anionic or neutral nature.⁵⁰ CPPs can be either protein-derived^{53,73} or synthetic ones.^{56,72} Other CPP classifications regard their amphipathic character and their conformation type (*e.g.* α -helical, β -sheet or other).^{50-52,75} CPPs may transport and deliver different types of bioactive molecules into living cells (anticancer drugs, RNA, DNA, or proteins).^{3-7,35-39,54,57-60} It is expected that CPPs derived from native

proteins and characterized by amphiphilic nature may be more efficient as drug delivery systems thanks to their self-assembly properties and low immunogenicity.

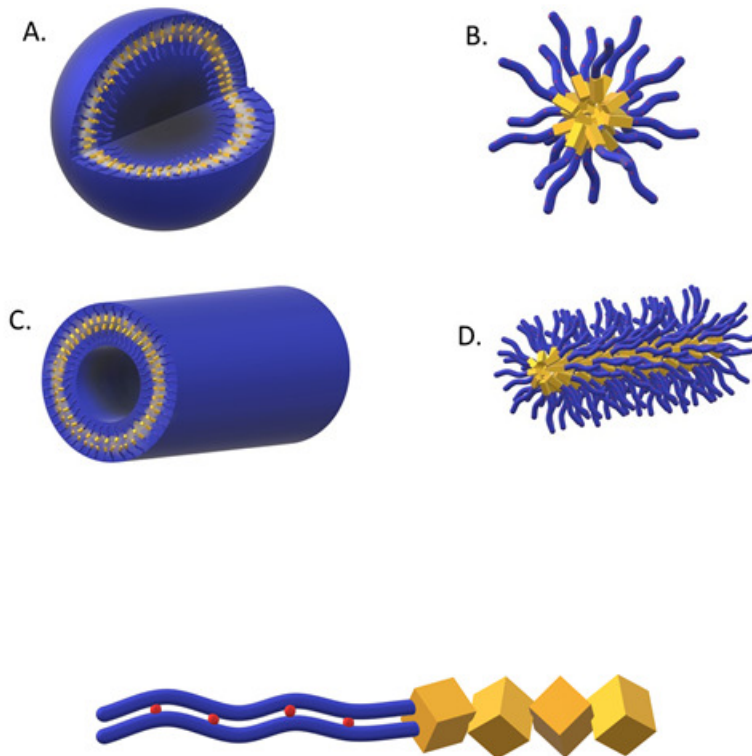


Figure 1. *Top panel:* Self-assembled nanostructures formed by amphiphilic peptides: (A) vesicles, (B) micelles, (C) nanotubes, and (D) rod-like objects exemplifying precursors of fibers. *Bottom panel:* Molecular architecture of an amphiphilic peptide composed of hydrophilic (blue) and hydrophobic (yellow) domains. The red dots indicate positively charged residues along the aminoacid sequence. The latter can be a cell penetrating peptide (CPP).

When immersed in aqueous media, amphiphilic molecules (lipids, surfactants, peptides, lipid-drug or peptide-drug conjugates, *etc.*) self-assemble through a spontaneous process leading to the formation of organized supramolecular structures.⁶⁴⁻⁸⁴ Figure 1 shows a few examples of nanoscale self-assembled architectures of amphiphilic derivatives. The underlying self-assembly mechanism is governed by hydrophobic interactions, dipole and electrostatic forces, π - π stacking, and hydrogen bonding.⁷⁵ In a first approximation, the shapes of the generated self-assembled aggregates are determined, according to the Israelachvili's concept, by the hydrophobic/hydrophilic balance and the sizes of the hydrophilic and hydrophobic domains of the molecules.^{66-68,75,76} However, a same chemical composition may yield different supramolecular aggregates depending on the environmental conditions (*e.g.* pH, salt type and concentration, temperature, *etc.*).^{30,33,70,77,80} The resulting topologies may comprise vesicles, spherical or elongated micelles, tubules and donut shapes of nanoscale dimensions as examples (Fig. 1). The shapes, which may globally span in space, correspond to indefinite nanotubes, fibers, and gels. The properties of the obtained amphiphilic nanostructures can be tuned by the inclusion of various functionalities towards the design of nano-biomaterials for a wide range of biomedical applications (*e.g.* vaccines, drug delivery, diagnostic imaging, tissue engineering and regeneration).^{4-7,26,27,33-39}

The present work exploits chosen computational approaches, developed on the basis of machine learning, for the design of amphiphilic peptides suitable for drug delivery applications. We first compare the published predictors of CPPs, some of which make use of artificial intelligence algorithms based on support vector machine (SVM) and random forest (RF) classifier methods.⁸⁵⁻¹⁰² We then apply selected web-accessible programs to the prediction of new peptide amphiphiles from the sequence of the human protein Ki-67 (as a protein of interest).

The combined results of the validated methods for CPP and transmembrane helix prediction will yield peptide sequences with natural biocompatibility and targeting properties, which are valuable characteristics for potential therapeutic applications. Based on such numerical computations, the research on amphiphilic peptides can be expanded to assist the design of safe and efficient nanocarriers for biomolecular delivery and specific targeting. In further steps, the stability of the predicted peptide scaffolds can be increased by the inclusion of unnatural amino acid residues. Various building blocks can be integrated towards functionalization of the peptides, for instance signal domains for improved selectivity, lipid anchors for enhanced amphiphilicity, vaccine epitopes, or fluorescent labels for theranostic purposes.^{3,13-15,26,27,37,84}

Methods

Choice between rational and machine learning assisted strategies for design of cell-penetrating peptides and amphiphilic derivatives thereof

The CPP class of peptides has gained significant interest since the finding that the HIV-1-derived Tat peptide sequence can cross the cellular membranes to trans-activate the viral promoter.^{61,62} Some of the first CPPs have been found serendipitously by studying the internalization of chemically synthesized peptides, which have been derived from the third helix of the homeodomain of *Antennapedia*.⁶³

Rational design of biomimetic cell-penetrating peptides has been performed by Karagiannis *et al.*,⁵⁵ Stupp *et al.*,^{76,78} and Zhang *et al.*⁷¹ among others.^{56,72,75,80} Peptide amphiphile structures consisting of up to four building blocks (a hydrophobic alkyl chain, a β -sheet forming segment, a bioactive epitope and charged groups) have been phenomenologically designed with the purpose to achieve both self-assembly and functional features of the generated constructs.⁷⁶ Molecular

dynamics (MD) simulations of self-assembled peptide systems have been performed by different groups.⁸² It has been experimentally shown that the hydrophobicity plays a crucial role for the shape of the resulting self-assembled aggregates.⁹ The structures produced with synthetic peptides, such as K₃-(WL)₃-W-NH₂, C₁₆-WAAAKAAAKAAAK, or C₁₈-KTTKS, have been in agreement with the molecular shape concept of the critical packing parameter, which determines the type of the obtained aggregates (micelles, vesicles or nanotube rods).^{66-68,71,72,83,84} Some studies have yielded ambiguous conclusions about the topology of the self-assembled peptide scaffolds because the Israelachvili's prediction does not account for the intermolecular interactions between α -helical moieties nor for the hydrogen bonding interactions between peptides of β -sheet conformations.^{4,13,34,75,80}

The initially proposed numerical approaches for CPP prediction have mainly been based on trial and error methods.^{57,58} These time-consuming analyses have permitted the constitution of a database of cell-penetrating peptides,^{46,85} but could not satisfy the growing market demand yet. The existing challenge has led to the emergence of an entirely new way for the prediction of amino acid sequences based on numerical methods.⁸⁵⁻¹⁰² The first generation of the *in silico* algorithms has been dedicated to CPPs predictions based on quantitative structure-activity relationship (QSAR) studies.⁸⁵ This method determines the score of a peptide according to the physico-chemical properties of its amino acid sequence. Such features have been calculated by Sandberg and colleagues through the principal component analysis (PCA) algorithm.⁸⁵ The considered descriptors, referred to as Z-factors, are related to the polarity, molecular weight, and orbital calculations. This approach has turned out to be simplistic because it has not taken into account the order of the amino acid residues in the studied peptide sequences. However, the obtained results have been essentially valuable as the method predicted about 95% of the known

CPPs to have cell-penetration characteristics. At the same time, only 68% of the known non-CPPs sequences have been predicted to be such. The limitation of this approach has been associated with the fact that the dataset used to train the algorithm has been composed of a small number of peptides. The latter have been extracted from heterogeneous experimental studies using different experimental parameters (concentrations, types of cells, *etc.*). Furthermore, it has been established that the training process of the QSAR method was biased, which may strongly affect the results of the predictions.

Subsequently, a new generation of algorithms came out with the development of artificial intelligence.⁸⁶⁻⁹² The first one has been based on the support vector machine classifier (SVM) and quantitative structure-activity relationship models.⁸⁶ The SVM involves information about the protein sequence, a position-specific aminoacid propensity, and physicochemical properties. The classifier has achieved a maximal accuracy of 83% by excluding the difficult aminoacid sequences. Sanders *et al.* have overcome this problem and have emphasized the importance of using balanced datasets of peptides.⁸⁷ The SVM classifier has been applied to a benchmark dataset using a total of 145 peptides, 111 of which were known CPPs and the others were known non-CPPs. Subsequently, the SVM classifier method has been used by Gautam *et al.* who have extended the benchmark dataset to a much larger set consisting of 708 peptides.⁸⁸ For this CellPPD predictor, the percentage of accuracy has increased to 81% on the independent datasets.

The machine training on neuronal network (NN) has emerged later and has become publicly available by the algorithm CPPpred conceived by Holton *et al.*⁸⁹ The accomplishment of this approach has revealed that the accuracy of the algorithm increases when the redundancy of the training sequences is reduced by 80% (despite of the small training set of CPPs).

Random Forest (RF) classifier is another machine learning approach to predict peptides with cell penetrating properties.⁹⁰ The RF algorithm integrates a number of decision trees and chooses the classification with the highest number of votes from the trees. The RF predictor is trained on various physicochemical and biochemical properties of amino acids and pairs of amino acids such as amino acid frequency, electrostatic charge, molecular volume, polarity, and secondary structure. In the random subsets of variables, each tree depends on the values of a random vector, which is sampled independently with the same distribution for all trees in the forest. A majority voting scheme is then applied for the classification purposes.

We have chosen the neural network (CPPpred) and the support vector machine classifier (CellPPD) algorithms for the prediction of novel CPP sequences from the nuclear protein Ki-67. These CPPs will be merged to hydrophobic signal sequences or membrane spanning fragments in order to form amphiphilic peptide scaffolds. The web-accessible predictor Skip-CPPpred, which is based on a random forest (RF) classification,⁹¹ will also be included in the employed here algorithm set.

Biological small-angle X-ray scattering (BioSAXS) for characterization of amphiphilic peptide nanoassemblies and particles

Biological small-angle X-ray scattering (BioSAXS), which comprises solution X-ray scattering, is an indispensable tool for the characterization of biological macromolecules and their propensity for self-assembly in aqueous media.¹⁰³⁻¹⁰⁸ Amphiphilic peptide assemblies represent randomly oriented supramolecular aggregates in a solution phase. The degree of arrangement and disorder leads to characteristic patterns when the samples are exposed to the X-ray beam. The scattering emerges upon interaction between the X-ray photons and the electrons

in the studied samples. Thus, the BioSAXS method provides ensemble-averaged results about the peptide molecules or particles dispersed in the aqueous medium.¹⁰³⁻¹⁰⁸

In aqueous environment, self-assembly of peptide amphiphiles into nanoscale aggregates occurs above a critical concentration. Structural evolution and transitions between different particle topologies are possible at increasing peptide concentrations. To probe the topologies of the peptide structural ensembles, the experimental SAXS profiles should be compared with computed theoretical scattering profiles. Such models have been developed for solutions of monodispersed identical particles with minimal interactions in a dispersed state.¹⁰⁹⁻¹¹² The derived analytical expressions of the scattering profiles provide information about the shape of the nanoscale objects through the form factors.¹⁰⁹

Particle form factors for various shape models can be considered in the analysis of the SAXS patterns, for instance monodispersed and homogeneous solid spheres, spherical shells, monodisperse cylinders, rods, discs, ellipsoids, cylinders with an elliptical cross section, or core/shell structures. The analytical models, employed in the particle shape analysis, describe the relationship between the scattered intensity and the particle shape for monodisperse particle solutions. Equations for form factors of different geometrical objects are used for this purpose.¹⁰⁹⁻¹¹¹

Exemplary SAXS plots of model functions, which simulate the scattering of nanoparticles of ideal geometrical shapes,¹¹² are presented in Figure 2 for selected values of the variable parameters.

(A) Spherical particles (solid spheres) are simulated by a model scattering function with a single parameter, *i.e.* the sphere radius (50 nm);

(B) Hollow spherical particles (vesicles) are described by a model scattering function with two parameters, namely the radius of the spherical shell (50 nm) and thickness of the lipid bilayer shell (3.5 nm);

(C) Hollow cylinders (nanotubes) are anisotropic particles, the scattering of which is simulated by three parameters, *i.e.* the length of the tube (300 nm), the radius of the cylindrical shell (50 nm) and the thickness of the lipid layer shell (3.5 nm);

(D) Nonflexible cylindrical rods (fibers) are simulated by two-parameter model scattering function – the length of the fiber (300 nm) and its radius (50 nm).

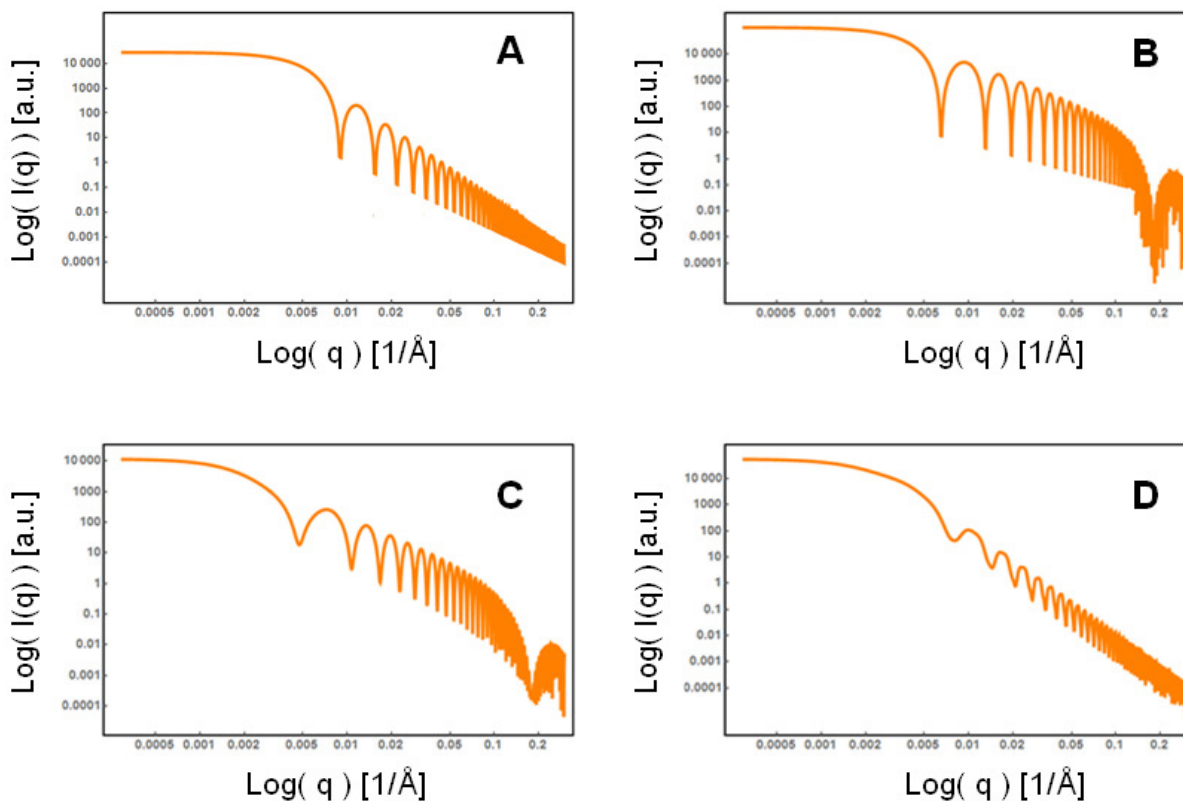


Figure 2. Examples of computed small-angle X-ray scattering (SAXS) patterns of monodispersed self-assembled amphiphilic peptide objects. The presented SAXS plots comprise

form factor scattering from a sphere (A), a hollow sphere (vesicle) (B), a hollow circular cylinder (nanotube) (C), and a rigid rod (fiber) (D).

BioSAXS experiments were performed at the SWING beamline of synchrotron SOLEIL (Saint Aubin, France). The samples were sealed in capillaries for X-ray measurements. The q -vector was defined as $q = (4\pi/\lambda) \sin\theta$, where 2θ is the scattering angle. The synchrotron radiation wavelength was $\lambda = 1.033 \text{ \AA}$. The synchrotron SAXS patterns were recorded with a two-dimensional EigerX 4-M detector (Dectris, Baden, Switzerland) at 12 keV. The exposure time was 1 s. No radiation damage was observed at this exposure time. An average of three spectra per capillary was acquired. Scattering patterns of an empty capillary and a capillary filled with MilliQ water were recorded for intensity background subtraction.

The small angle X-ray scattering (SAXS) data processing included fitting of the experimental scattering intensities by a library of analytical expressions available in the SASfit software package for SAXS analysis.¹¹²

Results

Design of new cell-penetrating peptides and application of the chosen set of algorithms for the prediction of amphiphilic aminoacid scaffolds from protein-derived sequences

The design of novel biocompatible CPP building blocks is inspired here by the features of the naturally occurring protein Ki-67.¹¹³⁻¹¹⁵ This human protein is composed of 3241 amino acids (see Fig. 3 below). It is an important nuclear protein that has been suggested to function as a

steric and electrostatic charge barrier against the aggregation of mitotic chromosomes.¹¹³ The cellular localization of the Ki-67 protein is a cell-cycle-phase dependent.

The performed peptide design takes into account that the spontaneous self-assembly process of protein-derived peptides is controlled by the hydrophobic/hydrophilic balance of the aminoacid constructs. The tendency for assembly into micelles, nanotubes, vesicles, nanofibers, nanobelts, or nanosheets in aqueous solutions will depend on the composition and the relative sizes of the two constituting building blocks (Fig. 1). The latter will impact the organization of the final aggregates. If the length of the hydrophobic domain is too long, the peptide solubility will be low. As a consequence, the corresponding critical aggregation concentration (CAC) might be too low for the development of drug delivery carriers. A sufficiently high concentration of positively charged amino acids in the hydrophilic domain can stabilize the thermodynamic equilibrium of the dispersed supramolecular aggregates (Fig. 1). Moreover, it can provide a higher affinity for interaction with the negatively charged cellular membranes, which is essential for efficient intracellular delivery of biomolecules.

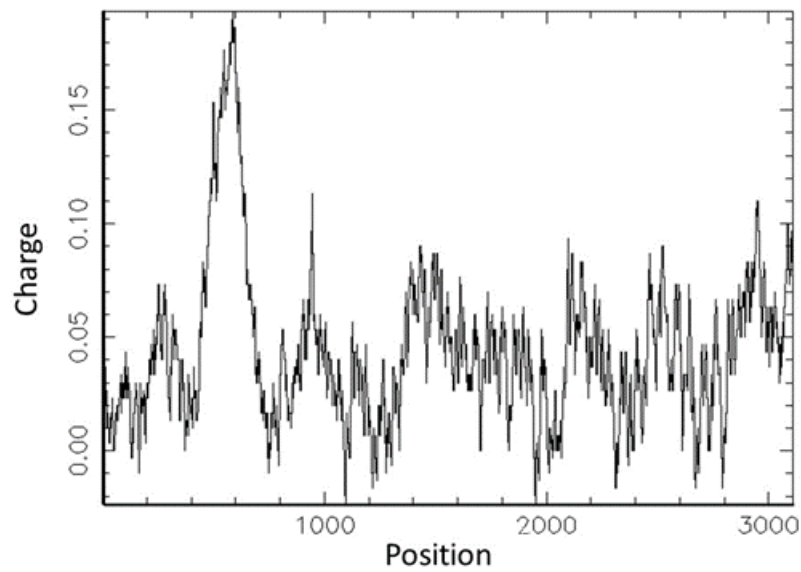
For the selection of candidate peptides with a membrane insertion potential, aminoacid sequences must be identified from protein domains that reside or interact with lipid plasma membranes. Transmembrane helices¹¹⁶ can be derived from the muscarinic acetylcholine receptor M2, the angiotensine 1A receptor, the metabotropic glutamate receptor, or the human galanin receptor type 1 as examples. Peptides with high membrane affinity involve also magainin 2, melittin, α -synuclein, orexin-A, lactoferricin, indolicidin, eledoisin, and endothelin-1 among others.^{10,58,74} The hydrophobic scaffold domains, derived from transmembrane spanning segments, will facilitate the membrane fusion, translocation, or endocytosis of the created peptides.

Our purpose here is to build, by the chosen computational tools, bio-inspired amphiphilic peptides comprising two distinct building blocks: a hydrophilic domain and a hydrophobic domain (Fig. 1, bottom panel). The protein-derived hydrophilic domain should involve positively charged amino acid fragments from the Ki-67 sequence, whereas the hydrophobic domain should be mainly composed of non-polar amino acids. The properties of the constructed protein-derived peptides will be computed and optimized by the cluster of web-accessible predictor software presented in Table 1.

Table 1. Machine learning based algorithms and web-based prediction tools employed for the calculation of the peptide properties.

Name	Principal	Outcomes
EMBOSS	Associates charges to amino acids and performs an average	Charge distribution of a protein sequence
I-Tasser	Hierarchical approach using different algorithms (Iterative Threading ASSEmblY Refinement)	Protein structure and function predictions; secondary structure prediction
CellPPD	Support vector machine (SVM)	CPP score
CPP pred	Neuronal Network (NN)	CPP score
SkipCPP-Pred	Random forest (RF) classifier	CPP score
TMHMM	Hidden Markov Model	Transmembrane helix prediction

Towards the identification of hydrophilic aminoacid sequences with cell penetrating properties, derived from the human antigen protein Ki-67, we started with the evaluation of its charge distribution using the EMBOSS system.¹¹⁷ Figure 3 (top panel) shows the localization of the largest positively charged aminoacid cluster identified along the sequence of the Ki-67 protein (Fig. 3, bottom panel).




```

1  MNPTRRLVITI KRSGVDGPHF PLSLSTCLFG RGIECDIRIQ LPVVSQCHCK IEIHEQEAIL
61  HNFSSNTPTQ VNGSVIDEPV RLKHGDVIT I IDRSFRYENE SLQNGRKS TE FPKKIRBQEP
121  ARRVSRSSEFS SDPDEKAQDS KAYSKITTEGK VSGNPNQVHIK NVKEDSTADD SKDSVVAQGT
181  NVHSSEHACR NCRNAADPI S GDFKEISSVK LVSRYGELKS VPTTQC LDNS KKNESPFQWL
241  YESVKKELDV KSKQENVLQY CRKSGLQTDY ATEKESADGL QGETQLLVSR KSRPKSGGSG
301  HAVAEPASPE QELDQNKGKQ RQVESVQTPS KAVGASFP LY EPAKMKTVPQ YSQQNSPQK
361  HKNKDLYTTG RRESVNLGKS EGFKAGDKTL TPKRLSTRNR TPAKVEDAAD SATKPENLSS
421  KTRGS IPTDV EVLPTETEIH NEFPLTLWLT QVERKIQKDS LS KPEKLGTT AGQMC SGLP
481  LSSVDINNFG DSINESEGIP LKRRRVSFGG HLRPELFDEN LPNPTPLKRG EAPT KRRSLV
541  MHTPPVLKKI IKEQPQPSGK QESGSEIHVE VKAQSILVIS P PAPS PRKTPV ASDQ KRRSCK
601  TAPASSSKSQ TEVPRKGRK SGNLPSKRV S ISRSQHDILQ MICS KRRSGA SEANLIVAKS
661  WADVVKLGAK QTQTKVIKHG PQRSMNKQR SPATPKKPVG EVHSQFSTGH ANSPCTIIG
721  KAHTKVVHP ARPVRVLNIF ISNQKDFKE DLSGIAEMFK TPKVQEQPQLT STCHIAISNS
781  ENLGRKQFQG TDSGEEP LLP TSE SFGGNVF FSAQNAARQP SDKCSAS PPL RRQCIRENGN
841  VAKTPRNTYK MTSLETKTSD TETEPSKTVS TANRSGRSTE FRNIQKLPVE SKSEETNTEI
901  VECILKRGQK ATLLQQRREG EMKEIERPFE TYKENIE LKE NDEKMKAMKR SRTWQKCAP
961  MSDLTDLKSL PDELMKQTA RQQLNLQTDQ HAKAPKSEK KIKTMPQSL QPEPINTPTH
1021  TKQQLKASLG KVGVKEE LLA VGFTRTSGE THTHRE PAG DGKSI RTFK SPQILDPA
1081  RVTGMKKWPR TPKEEAQSLE DLAGFKELFQ TPGPSEESMT DEKTKLACK SPPPEVDPT
1141  TSTKQWPKRS LRKADVEEF LALRKLTPSA GKAMLT PKPA GGDEKDIKAF MGTVPVKL
1201  AGTLPQSKRQ LQTPKEKQAQ LEDLAGFKEL FQTPGHTEEL VAAGKTKIP CDSPQSDPVD
1261  TPTSTKORPK RSIRKADVEG ELLACRNIMP SACKAMHTPK PSVGEEKDII IFVGTVPVKL
1321  DLTENLTGS KRPQTPKEEA QALEDLTGFK ELFQTPGHTE EAVAAGTKTK MPCESSPPES
1381  ADPTSTRRQ PKTPLEKRDV QKLSALKKL TQTSGETTHT DKVPGGEDKS INAFRETAQ
1441  KLDPAASVTG SKRHPKTKEK AQLLEDLAGL KE LFQTPVCT DKPTTHEKTT KIACRSQDPD
1501  VDTPTSSKQ SKRSLRKVDV EEEFFAL KKK TP SACKAMHT PKPAVSGEKN IYAFMGTPVQ
1561  KLDLTENLTG SKRELLQTPKE KQALEDLAG FKELFQTRGH TEESMTNDKT AKVACKSSQP
1621  DPDKNPASSK RRLKTSLGKV GVKEE LLAGV KLTQTSGETT HTHTEPTGDG KSMKAFMES
1681  KQILDASAASL TGSKRQLRTP KCKSEVPEDL AGFIELEFQTP SHTKESMTNE KTKPQSVYAS
1741  QPDLVDTPTS SKPQPKRSLR KADTEEEFLA FRHQTPSAGK AMHTPKPAVG EEK DINTFLG
1801  TPVQKLDQPG NLPQSNRRLLQ TRKKAQALE ELTGFERELFQ TPCDINPTTD ETKTKLILCK
1861  SPQSDPADTP TNTKORPKRS LRKADVEEF LAFRKLTPSA GKAMHTPKAA VGEK DINTF
1921  VGTPEVKLDDL LGNLPQS KRK PQTPEKAKA LEDLAGFKEL FQTPGHTEES MTDKITEVS
1981  CKSPQDPVK TPTSSKORLK ISLGKVGKVE EVLPVGLQTP TSCKTQTHR ETAGDGKSIK
2041  AFKESAKQML DPANYGTGME RWPRTPKEEA QLEDLAGFK ELFQTPDHTE ESTTDDKTK
2101  IACKSPPPES MDTPTSTRR PKTPLGKRDI VEELSALKQL TQTHHTDKVP QDEDKGINV
2161  RETAKCKLDP AASVTGSKRQ PRTPKCKAQP LEDLAGLKE L FQTPICTDKP TTHEKTKIA
2221  CRSPQDPVVG TPTIFKQSK RSLRKADVEE ES LALRKRTP SVGKAMDTPK PAGGDEKDMK
2281  AFMGTPVKL DLPGNLPQSK RWPQTPKEKA QALEDLAGE ELFQTPGTDK PTDEKTKI
2341  ACKSPQDPV DTPASTKORP KRNLRKADVE EEF LALRKRTP PSAGKAMDTPK KPAVSDEKNI
2401  NTFVETPVQK LLDLGNLPQS KRQQTPPEK AEALDLVGF KE LFQTPGHTE EESMTDDKIT
2461  EVSCKSPQPE SFKTSRSSKQ RLKIPLVKVD MKEEPLAVSK LTRTSGETTQ THTEPTGDSK
2521  SIKAFKESPK QILDPAASVT GSRROLRTRK EKARALEDLV DFKE LF SAGP HTEESMTIDK
2581  NTKIPCKSP PELTDTATST KRCPKTRPRK EVKEELSAVE RLQTSQST HTHKEPASGD
2641  EGIKVLKORA KKNPNVVEE PSRRRPRAP KKAQPLEDLA GFTE LSETSG HTQELTACK
2701  ATKIPCESPP LEVVDTTAST KRHLRTRVQK VQVKEEPSAV KFTQTSGETT DADKEPAGED
2761  KGIKALKESA KQTPAPAASV TCSRRRPRAP RESAQAIEDL AGFKDPAACH TEESMTDDKT
2821  TKIPCKSSPE LEDTATS SKR RPRTRAQKVE VKEE LLAGVK LQTSGETTH TDKEPVGEGK
2881  GTFKAFKQPAK RKLDAEDVIG SRRQPRAPKE KAQPLEDLAS FQELSQTPGH TEELANGAAD
2941  SFTSAPKQTP DSGKPLKISR KVLRAPKVEP VGDVSTRDP VKSQSKSNTS LPPLPKRGG
3001  GKDGSVTGTK RLRCMPAPEE IVEELPASKK QVAPRARCK SSEPVVIMKR SLRTSAKRIE
3061  PAEELNSNDM KTNKEEHLKQ DSVPENKGIS LRSRFRNKTE AEQITVEFV LAERIEINRN
3121  EKKPMKTSPE MDIQNPDDGA RKP IPRDKVT ENKRCILRSAR QNESSQPKVA EESGQKSAK
3181  VIMQNKKGK EAGNSDSMCL RSRKTKSQPA ASTLESKSVQ RVTRSVKCA ENPKKAEDNV
3241  CVKKIRTRSH RQSED

```

Figure 3. *Top panel:* Charge distribution along the amino acid sequence of the protein antigen Ki-67 represented by EMBOSS.¹¹⁷ The observed maximum is associated with the presence of a cluster (KRRR) of positively charged residues. The human protein Ki-67 (composed of 3241 amino acids) is encoded by the MKI67 gene. *Bottom panel:* Amino acid sequence of the human

Ki-67 protein with an indication of positively charged aminoacids (lysine [K] or arginine [R], yellow colour) and the clusters of charged residues (red and green colours).

Computational scans were realized with the chosen natural protein sequence (Fig. 3, bottom panel) using the set of algorithms listed in Table 1. This allowed identifying the regions of aminoacid residues that correspond to the positively charged fragments (as determined by the calculated charge distribution in Fig. 3). Then, a second scan of the algorithmic approach was realized in order to extract the potential cell-penetrating peptide (CPP) sequences inside the Ki-67 protein. Two algorithms were used for this purpose: CellCPP and CPPpred (Table 1). Based on the obtained scores, the most promising peptide fragments, each of 10aa length, were chosen for the constitution of the hydrophilic domain of the designed peptides.

Selected results of the performance of the machine learning prediction methods are presented in Table 2. The new cell penetrating peptides, which we derived from the biosurfactant protein Ki-67, are indicated in blue color as sequences of 10 amino acids (Table 2). Towards the objective to design novel amphiphilic peptides as candidate drug delivery carriers, we decided to merge the obtained new cell penetrating peptides (*e.g.* KKKARQKLVK) with chosen hydrophobic fragments, which may enable the targeting of the lipid plasma membranes or the intracellular membranes. Candidate hydrophobic chains of aminoacid residues were selected from the protein databank of transmembrane proteins (pdbtm).¹¹⁶ Interesting fragment sequences were chosen based on the criteria of simplicity, secondary structure, and length. Examples of such hydrophobic domain sequences, *e.g.* the mitochondria-targeting N-terminal region of the hexokinase-II protein and the membrane spanning segment of the acetylcholine receptor, are shown in Table 2 (yellow color).

Table 2. Examples of numerically predicted amphiphilic peptides and their properties calculated by the set of algorithms summarized in Table 1. The positively charged hydrophilic fragments are indicated in blue and the hydrophobic domains are indicated in yellow. Representative CPPpred and SkipCPP-pred scores are obtained for all cell penetrating peptides derived from the human protein Ki-67.

	Peptide amino acid sequence	Length (aa)	CPPpred score	SkipCPP-pred score	Transmembrane helix score
1:	KKKARQKLVKNLEFFYALLHSAIM	25	0.5	0.58	0
2:	GSRRRPRAPRMSTAISVLLAQAVFLLLSQ	29	0.6	0.87	0.6
3:	GSRRRPRAPT AISVLLAQAVFLLLSQ	27	0.75	0.89	0.2
4:	KRRPRTRAQKMSTAISVLLAQAVFLLLSQ	29	0.7	0.81	0.6

Practically, several amphipathic peptides were created by joining hydrophilic and hydrophobic moieties in all possible configurations. The resulting various combinations were numerically examined by the ensemble of algorithms (Table 1) for the prediction of their cell penetrating and transmembrane helix behaviour. The peptide chains were examined by I-Tasser (for the secondary structure and solvent accessibility), TMHMM (for the transmembrane probability rate and hydrophobicity), and CPPpred and SkipCPP-pred for the determination of the CPP scores (Table 2). The molecular models of the predicted peptide architectures with lengths of 25-29 residues long are shown in Figure 4.

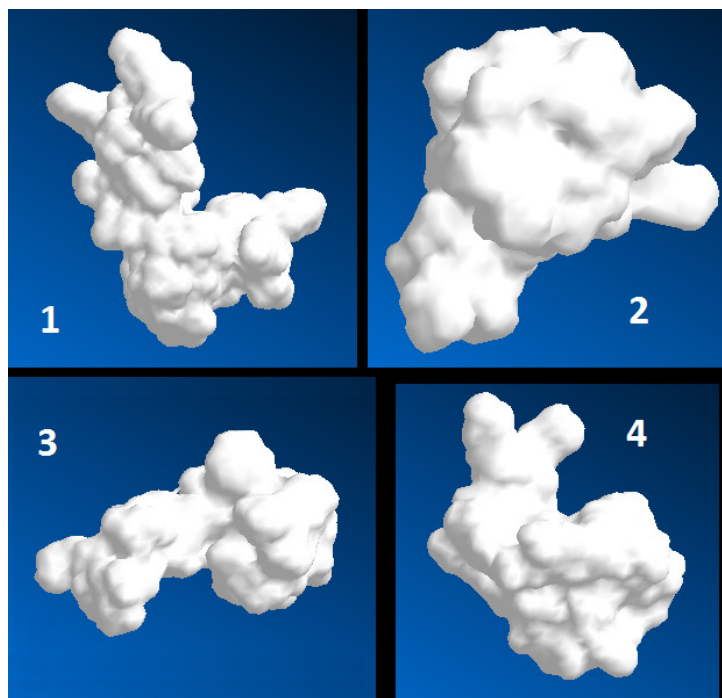


Figure 4. Molecular models of the numerically predicted amphiphilic peptide sequences given in Table 2. The cell penetrating peptide fragments are newly derived from the human nuclear protein Ki-67.

Discussion

Contribution of SAXS to establish the amphiphilic peptide self-assembly

Structural results obtained by SAXS revealed that the peptide amphiphile building blocks, predicted by the numerical approaches and yielding significant CPPpred and SkipCPP-pred scores, may easily self-assemble into nanoscale objects above a critical aggregation concentration. The SAXS data obtained from solutions of the predicted peptides were analyzed using analytical expressions restricted to simple shapes and geometries. An ellipsoidal shape of

the self-assembled aggregates was deduced from the analysis of the synchrotron small-angle scattering (BioSAXS) data shown in Figure 5. The performed fitting of the experimental scattering intensities by a model scattering function from a population of ellipsoids of revolution yielded the best outcome. The investigated amphiphilic peptide MIASHLLAYFFTELNKVLKQRAKKK has been designed with the purpose to target the mitochondria. Liu *et al.* have shown that the critical aggregation concentration of the peptide can be decreased by anchoring of a long fatty acid chain to the hydrophobic terminus of the aminoacid sequence MIASHLLAYFFTELN.⁹ This effect has enhanced the bioactivity of peptide conjugate upon mitochondrial targeting of lung cancer cells.⁹

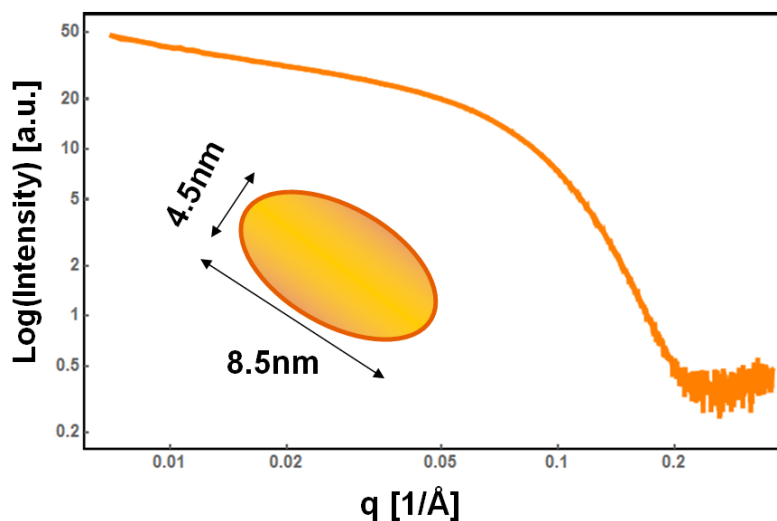


Figure 5. Synchrotron small-angle scattering (BioSAXS) pattern presenting an experimental evidence for the self-assembly propensity of the predicted amphiphilic peptide MIASHLLAYFFTELNKVLKQRAKKK in an aqueous medium. The peptide construct consists of the mitochondria-targeting N-terminal region of the protein hexokinase-II

(MIASHLLAYFFTELN) and a cell penetrating peptide derived from the Ki-67 sequence (*i.e.* KVLKQRAKKK).

The shapes and the sizes of the self-assembled amphiphilic peptide aggregates are crucial for the drug delivery efficacy in addition to their charge and interfacial characteristics. The transport and diffusion properties of nanocarriers with elongated shapes are different from those of the sphere-shaped particles. The latter move faster and accumulate to a lower extent at interfaces. At variance, the circulation time of the rod-like particles is prolonged with regards to spherical ones. Thus, the nanocarrier shape and geometry have impact on the drug transport across the biological barriers and the cellular internalization process.

Elongated, rod- and tubule-shaped nanoparticles may be more efficient in controlled release applications because of the increased circulation time and the promoted adhesive interactions with the cellular membranes. However, it is challenging to achieve a control over their asymmetric geometries in the nanoscale range. Moreover, such long objects may cause remodeling of the plasma membranes.

Quite often, the self-assembly of amphiphilic peptides yields fibers, which considerably exceed nanoscale dimensions and form micrometer-long aggregates.^{7,76} Long nanotubes may also reach macroscopic sizes. Delivery carriers of such micrometer-large dimensions are challenging for a systemic administration of drugs and generally should be avoided for parenteral use (via injections). Consequently, topical applications, such as implants and intracutaneous administration, have been envisioned for their clinical use.⁷

Drug delivery carriers of sizes typically below 100 nm can evade the systemic clearance. Therefore, the ellipsoid-like nanoassemblies, as those reported in Figure 5, may be expected to display higher therapeutic efficacy in parenteral drug delivery.

Discussion on the applications of amphiphilic peptides in drug delivery

The advantages of using peptides as delivery systems rather than other chemicals rely on their biocompatibility, their low cost and facile synthesis using standard solid-phase techniques, their tunable bioactivity, and targeting opportunities including for instance nuclear localization sequences (NLS) or mitochondria targeting fragments. The peptide self-assembly into nanostructures can provide reservoirs with enhanced drug delivery efficiency and can permit the extension of the half-life of the encapsulated biomolecules. Peptide-based drug delivery carriers comprise various supramolecular architectures including nanotubes, vesicles, micelles, cubosomes, and hydrogels constituted of self-assembled fibers.^{4,6,7,13,15,24,26,33-35} Vesicles and micelles offer interesting possibilities to encapsulate or covalently bind therapeutic molecules.^{20,22,26,27,76-78}

The research on application of amphiphilic peptides in therapeutic delivery and regenerative medicine (*e.g.* ocular drug delivery, vaccine delivery, and anticancer agents delivery) constantly increases.^{3-10,13-24,35-39} Besides the wide range of potential biomedical applications in tissue engineering, imaging and diagnostics, comprehensive reviews have been published about the use of amphiphilic peptides for delivery of biotherapeutics such as oligonucleotides, proteins, or siRNA.^{3,7,23,24,26,35,37,51,52,58,75-84} Cell-penetrating peptide amphiphiles can form complexes with siRNAs and mediate the delivery of nucleic acids toward gene silencing in gene therapy.^{35,53-58} They have successfully delivered drugs and photosensitisers *in vitro* and *in vivo*, for instance in

cultured rabbit lens epithelial cells.¹⁹ The reported uptake mechanisms are varied among endocytosis, membrane breaking and translocation, or macropinocytosis.

One of the most promising research directions on peptide amphiphiles includes cancer treatment strategies.^{5,20-22} Recent reports have shown that treatments of breast cancer could be improved by coupling the active drug molecule to a cell-penetrating peptide.^{20,37,39} The conjugation of an anti-cancer cargo to a CPP can improve tumor targeting and the *in vivo* drug transport, thus reducing the side effects of the therapy.^{20,39}

Peptide-based hydrogels have been used for controlled drug release and are the most common and studied self-assembled peptide nanostructures at the time.^{6,7,67,78,84} The regenerative potential of hydrogels composed of peptide amphiphile nanofibers has been examined in various applications, *e.g.* traumatic injury of spinal cord, bone regeneration, and promotion of neurite formation.^{67,79} Peptide amphiphile hydrogels have provided sustained release of therapeutic proteins (*e.g.* trophic factors) in topical applications.⁷ The bioactivity of the entrapped proteins has been preserved over long periods. Hydrogels of supramolecular peptide amphiphile nanofibers, loaded with drugs and high affinity antibodies, can suppress tumor growth in cancer therapy. A concrete example is the RADA16-I peptide,¹⁵ which has been the subject of a number of studies on protein release kinetics, breast tissue reconstruction, and also for the elaboration of peptide fibrous scaffolds for commercial purposes (*e.g.* the *PuraMatrix*TM product).

Conclusion

Taking into account the expanding applications of amphiphilic peptides in nanobiomaterials science and regenerative nanomedicine, the development of efficient algorithms for CPPs

prediction and evaluation of the peptide properties is crucial for the acceleration of the translational research into clinics.

We proposed a combined computational design approach, which comprises a set of algorithms that is applied for scanning of naturally occurring protein sequences for CPP fragments towards the identification of candidate amphiphilic peptides with desired features. The amphiphilic peptide scaffolds are created from at least two building blocks. The hydrophobic domains can be designed from transmembrane protein fragments, which are available in the protein databank of transmembrane proteins (pdbtm). The hydrophilic domains can be designed as CPP sequences, which are derived from proteins with biosurfactant properties. In perspective, the outcome of the suggested combined algorithm may include also the uptake score of the CPPs and their mode of cellular internalization.

ACKNOWLEDGMENT

B.A. acknowledges the supports from the projects “Structural dynamics of biomolecular systems” (ELIBIO) (CZ.02.1.01/0.0/0.0/15_003/0000447) and “Advanced research using high-intensity laser produced photons and particles” (CZ.02.1.01/0.0/0.0/16_019/0000789) from the European Regional Development Fund. We acknowledge the allocation of beam time at Synchrotron SOLEIL (France) through the projects 20170933 and 20181489 and the scientific and technical support at the SWING beamline.

AUTHOR INFORMATION

Corresponding Author

*e-mail: Angelina.Angelova@universite-paris-saclay.fr (A.A.).

ORCID

Angelina Angelova: 0000-0002-0285-0637

Borislav Angelov: 0000-0003-3131-4822

REFERENCES

1. Hassanzadeh, P.; Atyabi, F.; Dinarvand R. The Significance of Artificial Intelligence in Drug Delivery System Design. *Adv. Drug Delivery Rev.* **2019**, *151-152*, 169-190.
2. Vamathevan, J.; Clark, D.; Czodrowski, P.; Dunham, I.; Ferran, E.; Lee, G.; Li, B.; Madabhushi, A.; Shah, P.; Spitzer M.; Zhao, S. Applications of Machine Learning in Drug Discovery and Development. *Nature Reviews Drug Discovery.* **2019**, *18*, 463–477.
3. Tesauro, D.; Accardo, A.; Diaferia, C.; Milano, V.; Guillon, J.; Ronga, L.; Rossi, F. Peptide-Based Drug-Delivery Systems in Biotechnological Applications: Recent Advances and Perspectives. *Molecules*, **2019**, *24*(2), 351.
4. Wang, Q.; Zhang, X.; Zheng, J.; Liu, D. Self-Assembled Peptide Nanotubes as Potential Nanocarriers for Drug Delivery. *RSC Adv.*, **2014**, *4*, 25461-25469.
5. Liu, K.; Xing, R.; Zou, Q.; Ma, G.; Mohwald, H.; Yan, X. Simple Peptide-Tuned Self-Assembly of Photosensitizers towards Anticancer Photodynamic Therapy. *Angew. Chem., Int. Ed.* **2016**, *55*, 3036– 3039.
6. Nagai, Y.; Unsworth, L.D.; Koutsopoulos, S.; Zhang, S. Slow Release of Molecules in Self-Assembling Peptide Nanofiber Scaffold. *J. Controlled Rel.* **2016**, *115*, 18–25.
7. Sato, K.; Hendricks, M.P.; Palmer, L.C. Stupp, S.I. Peptide Supramolecular Materials for Therapeutics. *Chem. Soc. Rev.* **2018**, *47*, 7539-7551.
8. Angelova, A.; Drechsler, M.; Garamus, V.M.; Angelov, B., Pep-lipid Cubosomes and Vesicles Compartmentalized by Micelles from Self-Assembly of Multiple Neuroprotective Building Blocks Including a Large Peptide Hormone PACAP-DHA. *ChemNanoMat*, **2019**, *5*, 1381–1389.
9. Liu, D.; Angelova, A.; Liu, J.; Garamus, V.M.; Angelov, B., Zhang, X.; Li, Y.; Feger, G.; Li, N.; Zou, A. Self-Assembly of Mitochondria-Specific Peptide Amphiphiles Amplifying Lung Cancer Cell Death through Targeting the VDAC1–Hexokinase-II Complex. *J. Mater. Chem. B*, **2019**, *7*, 4706-4716.
10. Song, Z.; Chen, X.; You, X.; Huang, K.; Dhinakar, A.; Gu, Z.; Wu, J. Self-Assembly of Peptide Amphiphiles for Drug Delivery: The Role of Peptide Primary and Secondary Structures. *Biomater Sci.* **2017**, *5*, 2369-2380.

11. Angelova, A.; Angelov, B. Dual and Multi-Drug Delivery Nanoparticles Towards Neuronal Survival and Synaptic Repair. *Neural Regeneration Research*, **2017**, *12*, 886-889.
12. Angelova, A.; Angelov, B.; Mutafchieva, R.; Lesieur, S.; Couvreur, P. Self-Assembled Multicompartment Liquid Crystalline Lipid Carriers for Protein, Peptide, and Nucleic Acid Drug Delivery. *Accounts of Chemical Research* **2011**, *44*, 147-156.
13. Wang, Y.; Zhang, Y.; Li, X.; Li, C.; Yang, Z.; Wang, L. A Peptide-Based Supramolecular Hydrogel for Controlled Delivery of Amine Drugs. *Chemistry Asian J.*, **2018**, *13*, 3460-3463.
14. Van Hell, A.J.; Fretz, M.M.; Crommelin, D.J.; Hennink, W.E.; Mastrobattista, E. Peptide Nanocarriers for Intracellular Delivery of Photosensitizers. *Journal of Controlled Release*, **2010**, *141*, 347-353.
15. Wu, H.; Zhou, T.; Tian, L.; Xia, Z.; Xu, F. Self-Assembling RADA16-I Peptide Hydrogel Scaffold Loaded with Tamoxifen for Breast Reconstruction. *BioMed Research International*, **2017**, article ID 3656193.
16. Hamley, I.W. Small Bioactive Peptides for Biomaterials Design and Therapeutics. *Chem. Rev.* **2017**, *117*, 14015–14041.
17. Liu, Q.; Wang, J.; Dong, Y.D.; Boyd, B.J. Using a Selective Cadmium-Binding Peplipid to Create Responsive Liquid Crystalline Nanomaterials. *J. Colloid Interface Sci.*, **2015**, *449*, 122-129.
18. Ding, J.; Li, C.; Zhang, Y.; Xu, W.; Wang, J.; Chen, X. Chirality-Mediated Polypeptide Micelles for Regulated Drug Delivery. *Acta Biomaterialia*, **2015**, *11*, 346-355.
19. Pescina, S.; Ostacolo, C.; Gomez-Monterrey, I.M.; Sala, M.; Bertamino, A.; Sonvico, F.; Padula, C.; Santi, P.; Bianchera, A.; Nicoli, S. Cell Penetrating Peptides in Ocular Drug Delivery: State of the Art. *Journal of Controlled Rel.* **2018**, *284*, 84-102.
20. Aroui, S.; Brahim, S.; De Waard, M.; Kenani, A. Cytotoxicity, Intracellular Distribution and Uptake of Doxorubicin and Doxorubicin Coupled to Cell-Penetrating Peptides in Different Cell Lines: A Comparative Study. *Biochem. Biophys. Res. Commun.* **2010**, *391*, 419–425.
21. Nakase, I.; Konishi, Y.; Ueda, M.; Saji, H.; Futaki, S. Accumulation of Arginine-Rich Cell-Penetrating Peptides in Tumors and the Potential for Anticancer Drug Delivery in Vivo. *Journal of Controlled Release* **2012**, *159*, 181–188.
22. Al-Ahmady, Z.S.; Al-Jamal, W.T.; Bossche, J.V.; Bui, T.T.; Drake, A.F.; Mason, A.J.; Kostarelos, K. Lipid-Peptide Vesicle Nanoscale Hybrids for Triggered Drug Release by Mild Hyperthermia in Vitro and in Vivo. *ACS Nano*. **2012**, *6*, 9335-46.
23. Qiu, F.; Chen, Y.; Tang, C.; Zhao, X. Amphiphilic Peptides as Novel Nanomaterials: Design, Self-Assembly and Application. *Int J Nanomedicine*, **2018**, *13*, 5003-5022.
24. Rad-Malekshahi, M.; Lempsink, L.; Amidi, M.; Hennink, W.E.; Mastrobattista, E. Biomedical Applications of Self-Assembling Peptides. *Bioconjugate Chem.* **2016**, *27*, 3-18.
25. Fong, W.K.; Negrini, R.; Vallooran, J.J.; Mezzenga, R.; Boyd, B.J. Responsive Self-Assembled Nanostructured Lipid Systems for Drug Delivery and Diagnostics. *J. Colloid Interface Sci.* **2016**, *484*, 320-339.
26. Eskandari, S.; Guerin, T.; Toth, I.; Stephenson, R.J. Recent Advances in Self-Assembled Peptides: Implications for Targeted Drug Delivery and Vaccine Engineering. *Adv Drug Deliv Rev.* **2017**, *110-111*, 169-187.

27. Fatouros, D.G.; Lamprou, D.A.; Urquhart, A.J.; Yannopoulos, S.N.; Vizirianakis, I.S.; Zhang, S.; Koutsopoulos, S. Lipid-Like Self-Assembling Peptide Nanovesicles for Drug Delivery, *ACS Appl. Mater. Interfaces*, **2014**, *6*, 8184–8189.
28. Angelova, A.; Fajolles, C.; Hocquelet, C.; Djedaïni-Pilard, F.; Lesieur, S.; Bonnet, V.; Perly, B.; Lebas, G.; Mauclaire, L. Physico-Chemical Investigation of Asymmetrical Peptidolipidyl-Cyclodextrins, *Journal of Colloid and Interface Science*, **2008**, *322*, 304-314.
29. Shao, X.; Bor, G.; Al-Hosayni, S.; Salentinig, S.; Yaghmur, A. Structural Characterization of Self-Assemblies of New Omega-3 Lipids: Docosahexaenoic Acid and Docosapentaenoic Acid Monoglycerides, *Phys. Chem. Chem. Phys.* **2018**, *20*, 23928-23941.
30. Tajima, K.; Koshinuma, M.; Nakamura, A.; Gershfeld, N.L. Sponge-Vesicle Transformation in Binary Mixtures of Ionized Phospholipid Bilayers. *Langmuir*, **2000**, *16*, 2576–2580.
31. Angelova, A.; Angelov, B.; Garamus, V.M.; Drechsler, M. A Vesicle-to-Sponge Transition via the Proliferation of Membrane-Linking Pores in Omega-3 Polyunsaturated Fatty Acid-Containing Lipid Assemblies. *J. Mol. Liquids*, **2019**, *279*, 518-523.
32. Chong, J.Y.T.; Mulet, X.; Waddington, L.J.; Boyd B.; Drummond, C.J. Steric Stabilisation of Self-Assembled Cubic Lyotropic Liquid Crystalline Nanoparticles: High throughput Evaluation of Triblock Polyethylene Oxide-Polypropylene Oxide-Polyethylene Oxide copolymers. *Soft Matter*, **2011**, *7*, 4768–4777.
33. Pandit, G.; Roy, K.; Agarwal, U.; Chatterjee, S. Self-Assembly Mechanism of a Peptide-Based Drug Delivery Vehicle. *ACS Omega*, **2018**, *3*, 3143–3155.
34. Accardo, A.; Morisco, A.; Palladino, P.; Palumbo, R.; Tesauro, D.; Morelli, G. Amphiphilic CCK Peptides Assembled in Supramolecular Aggregates: Structural Investigations and In Vitro Studies. *Mol. Biosyst.* **2011**, *7*, 862–870.
35. Accardo, A.; Morisco, A.; Tesauro, D.; Pedone, C.; Morelli, G. Naposomes: A New Class of Peptide-Derivatized, Target-Selective Multimodal Nanoparticles for Imaging and Therapeutic Applications. *Ther. Deliv.* **2011**, *2*, 235–257.
36. Pujals, S.; Fernandez-Carneado, J.; Lopez-Iglesias, C.; Kogan, M.J.; Giralt, E. Mechanistic Aspects of Cell-Penetrating Peptide-Mediated Intracellular Drug Delivery: Relevance of CPP Self-Assembly. *Biochim. Biophys. Acta Biomembranes*, **2006**, *1758*, 264–279.
37. Yang, Y.; Yang, Y.; Xie, X.; Cai, X.; Mei, X. Preparation and Characterization of Photo-Responsive Cell-Penetrating Peptide-Mediated Nanostructured Lipid Carrier. *J. Drug Target.* **2014**, *22*, 891–900.
38. Raymond, D. M.; Nilsson, B. L. Multicomponent Peptide Assemblies. *Chem. Soc. Rev.*, **2018**, *47*, 3659-3720.
39. Jiang, Y.; Pang, X.; Liu, R.; Xiao, Q.; Wang, P.; Leung, A.W.; Luan, Y.; Xu, C. Design of an Amphiphilic RGD Peptide and Self-Assembling Nanovesicles for Improving Tumor Accumulation and Penetration and the Photodynamic Efficacy of the Photosensitizer. *ACS Appl. Mater. Interfaces* **2018**, *10*, 31674–31685.
40. Zhavoronkov, A. Artificial Intelligence for Drug Discovery, Biomarker Development, and Generation of Novel Chemistry. *Molecular Pharmaceutics* **2018**, *15*, 4311-4313.
41. Giguère, S.; Laviolette, F.; Marchand, M.; Tremblay, D.; Moineau, S.; Liang, X.; Biron, E.; Corbeil, J. Machine Learning Assisted Design of Highly Active Peptides for Drug Discovery. *PLoS Comput Biol.* **2015**, *11*, pe1004074.

42. Damiati, S.A.; Alaofi, A.L.; Dhar, P.; Alhakamy, N.A. Novel Machine Learning Application for Prediction of Membrane Insertion Potential of Cell-Penetrating Peptides. *Int. J Pharm.* **2019**, *567*, 118453.
43. Wolfe, J.M.; Fadzen, C.M.; Choo, Z.N.; Holden, R.L.; Yao, M.; Hanson, G.J.; Pentelute, B.L. Machine Learning To Predict Cell-Penetrating Peptides for Antisense Delivery. *ACS Central Sci.*, **2018**, *4*, 512-520.
44. Regberg, J.; Srimanee, A.; Erlandsson, M.; Sillard, R.; Dobchev, D.A.; Karelson, M.; Langel, U. Rational Design of a Series of Novel Amphipathic Cell-Penetrating Peptides. *Int J Pharm.* **2014**, *464*, 111-116.
45. Hallbrink, M.; Kilk, K.; Elmquist, A.; Lundberg, P.; Lindgren, M.; Jiang, Y.; Pooga, M.; Soomets, U.; Langel, Ü. Prediction of Cell-Penetrating Peptides. *Int. J. Pept Res. Ther.* **2005**, *11*, 249–259.
46. Tang, J.; Ning, J.; Liu, X.; Wu, B.; Hu, R. A Novel Amino Acid Sequence-Based Computational Approach to Predicting Cell-penetrating Peptides. *Curr. Comput Aided Drug Des.*, **2019**, *15*, 206-211.
47. Manavalan, B.; Subramaniyam, S.; Shin, T.H.; Kim, M.O.; Lee, G. Machine-Learning-Based Prediction of Cell-Penetrating Peptides and Their Uptake Efficiency with Improved Accuracy. *J. Proteome Res.*, **2018**, *17*, 2715-2726.
48. Wakabayashi, R.; Suehiro, A.; Goto, M.; Kamiya, N. Designer Aromatic Peptide Amphiphiles for Self-Assembly and Enzymatic Display of Proteins with Morphology Control. *Chem. Commun.* **2019**, *55*, 640-643.
49. Basith, S.; Manavalan, B.; Hwan Shin, T.; Lee, G. Machine Intelligence in Peptide Therapeutics: A Next-Generation Tool for Rapid Disease Screening. *Med. Res. Rev.* **2020**, *16* DOI: 10.1002/med.21658.
50. Milletti, F. Cell-Penetrating Peptides: Classes, Origin, and Current Landscape, *Drug Discovery Today*, **2012**, *17*, 850-860.
51. Lindgren, M.; Hällbrink, M.; Prochiantz, A.; Langel, U. Cell-Penetrating Peptides, *Trends Pharmacol. Sci.*, **2000**, *21*, 99-103.
52. Copolovici, D.M.; Langel, K.; Eriste, E.; Langel, U. Cell-Penetrating Peptides: Design, Synthesis, and Applications. *ACS Nano*, **2014**, *8*, 1972–1994.
53. Liu, B.R.; Huang, Y.-W.; Aronstam, R.S.; Lee, H.-J. Identification of a Short Cell-Penetrating Peptide from Bovine Lactoferricin for Intracellular Delivery of DNA in Human A549 Cells. *PLoS One*, **2016**, *11(3)*, e0150439.
54. McClorey, G.; Banerjee, S. Cell-Penetrating Peptides to Enhance Delivery of Oligonucleotide-Based Therapeutics. *Biomedicines* **2018**, *6(2)*, 51.
55. Karagiannis, E.D.; Urbanska, A.M.; Sahay, G.; Pelet, J.M.; Jhunjhunwala, S.; Langer, R.; Anderson, D.G. Rational Design of a Biomimetic Cell Penetrating Peptide Library. *ACS Nano* **2013**, *7*, 8616–8626.
56. Zhao, X. Design of Self-Assembling Surfactant-Like Peptides and their Applications. *Curr. Opin. Colloid & Interface Sci.* **2009**, *14*, 340-348.
57. Wender, P.A.; Mitchell, D.J.; Pattabiraman, K.; Pelkey, E.T.; Steinman, L.; Rothbard, J.B. The Design, Synthesis, and Evaluation of Molecules that Enable or Enhance Cellular Uptake: Peptoid Molecular Transporters. *Proc Natl Acad Sci USA.* **2000**, *97*, 13003-8.

58. Futaki, S.; Suzuki, T.; Ohashi, W.; Yagami, T.; Tanaka, S.; Ueda, K.; Sugiura, Y. Arginine-Rich Peptides. An Abundant Source of Membrane-Permeable Peptides Having Potential as Carriers for Intracellular Protein Delivery. *J. Biol. Chem.* **2001**, *276*, 5836-40.
59. Pirisinu, M.; Blasco, P.; Tian, X.; Sen, Y.; Bode, A.M.; Liu, K.; Dong, Z. Analysis of Hydrophobic and Hydrophilic Moments of Short Penetrating Peptides for Enhancing Mitochondrial Localization: Prediction and Validation. *FASEB J.* **2019**, *33*, 7970-7984.
60. Jain, A.; Chugh, A. Mitochondrial Transit Peptide Exhibits Cell Penetration Ability and Efficiently Delivers Macromolecules to Mitochondria. *FEBS Lett.* **2016**, *590*, 2896-905.
61. Frankel, A.D.; Pabo, C.O. Cellular Uptake of the Tat Protein from Human Immunodeficiency Virus. *Cell*, **1988**, *55*, 1189–1193.
62. Green, M.; Loewenstein, P.M. Autonomous Functional Domains of Chemically Synthesized Human Immunodeficiency Virus Tat Trans-Activator Protein. *Cell* **1988**, *55*, 1179–1188.
63. Derossi, D.; Joliot, A.H.; Chassaing, G.; Prochiantz, A. The Third Helix of the *Antennapedia* Homeodomain Translocates through Biological Membranes. *J. Biol. Chem.* **1994**, *269*, 10444-50.
64. Mikhalevich, V.; Craciun, I.; Kyropoulou, M.; Palivan, C.G.; Meier, W. Amphiphilic Peptide Self-Assembly: Expansion to Hybrid Materials. *Biomacromolecules*, **2017**, *18*, 3471-3480.
65. Löwik, D.W., van Hest, J.C. Peptide Based Amphiphiles. *Chem Soc Rev.* **2004**, *33*, 234-45.
66. Edwards-Gayle, C.J.C., Hamley, I.W. Self-Assembly of Bioactive Peptides, Peptide Conjugates, and Peptide Mimetic Materials. *Org. Biomol. Chem.* **2017**, *15*, 5867–5876.
67. Dehsorkhi, A.; Castelletto, V.; Hamley, I.W. Self-Assembling Amphiphilic Peptides. *J. Pept Sci.*, **2014**, *20*, 453-67.
68. Dasgupta, A.; Das, D. Designer Peptide Amphiphiles: Self-Assembly to Applications. *Langmuir*, **2019**, *35*, 10704-10724.
69. Chen, J.; Zou X. Self-Assemble Peptide Biomaterials and Their Biomedical Applications. *Bioact Mater.* **2019**, *4*, 120-131.
70. Imura, T.; Tsukui, Y.; Taira, T.; Aburai, K., Sakai, K.; Sakai, H., Abe M., Kitamoto, D. Surfactant-like Properties of an Amphiphilic α -Helical Peptide Leading to Lipid Nanodisc Formation. *Langmuir* **2014**, *30*, 4752-4759.
71. Santoso, S.; Hwang, W.; Hartman, H.; Zhang, S. Self-Assembly of Surfactant-Like Peptides with Variable Glycine Tails to Form Nanotubes and Nanovesicles. *Nano Lett.* **2002**, *2*, 687–691.
72. Thota, N.; Jiang, J. Self-Assembly of Amphiphilic Peptide (AF)6H5K15 Derivatives: Roles of Hydrophilic and Hydrophobic Residues. *J. Phys. Chem. B* **2014**, *118*, 2683-2692.
73. Elaid, S.; Libersou, S.; Ouldali, M.; Morellet, N.; Desbat, B.; Alves, I.D.; Lepault, J.; Bouaziz, S. A Peptide Derived from the Rotavirus Outer Capsid Protein VP7 Permeabilizes Artificial Membranes. *Biochim Biophys Acta* **2014**, *1838*, 2026-35.
74. Almeida, P.F. Membrane-Active Peptides: Binding, Translocation, and Flux in Lipid Vesicles. *Biochim Biophys Acta*, **2014**, *1838*, 2216-27.
75. Wang, J.; Liu, K.; Xing, R.; Yan, X. Peptide Self-Assembly: Thermodynamics and Kinetics. *Chem. Soc. Rev.*, **2016**, *45*, 5589-5604.
76. Hendricks, M.P.; Sato, K.; Palmer, L.C.; Stupp, S.I. Supramolecular Assembly of Peptide Amphiphiles. *Acc Chem Res.* **2017**, *50*, 2440-2448.
77. Kimura, S.; Kim, D-H.; Sugiyama, J.; Imanishi, Y. Vesicular Self-Assembly of a Helical Peptide in Water, *Langmuir*, **1999**, *15*, 4461-4463.

78. Cui, H.; Muraoka, T.; Cheetham, A.G.; Stupp, S.I. Self-Assembly of Giant Peptide Nanobelts. *Nano Lett.*, **2009**, *9*, 945–951.
79. San, B.H.; Hwang, J.; Sampath, S.; Li, Y.; Bennink, L.L.; Yu, S.M. Self-Assembled Water-Soluble Nanofibers Displaying Collagen Hybridizing Peptides. *J. Am. Chem. Soc.* **2017**, *139*, 16640-16649.
80. Zhao, Y.; Deng, L.; Wang, J.; Xu, H.; Lu, J.R. Solvent Controlled Structural Transition of KI4K Self-Assemblies: From Nanotubes to Nanofibrils. *Langmuir*, **2015**, *31*, 12975-12983.
81. Makam, P.; Gazit, E. Minimalistic Peptide Supramolecular Co-Assembly: Expanding the Conformational Space for Nanotechnology. *Chem. Soc. Rev.*, **2018**, *47*, 3406-3420.
82. Gudlur, S.; Sukthankar, P.; Gao, J.; Avila, L.A.; Hiromasa, Y.; Chen, J.; Iwamoto, T.; Tomich, J.M. Peptide Nanovesicles Formed by the Self-Assembly of Branched Amphiphilic Peptides. *PLoS One*, **2012**, *7*, e45374.
83. Sigg, S.J.; Schuster, T.B.; Meier, W.P. Self-Assembled Structures from Amphiphilic Peptides. *Chimia*, **2013**, *67*, 881–884.
84. Palladino, P.; Castelletto, V.; Dehsorkhi, A.; Stetsenko, D.; Hamley, I.W. Conformation and Self-Association of Peptide Amphiphiles Based on the KTTKS Collagen Sequence. *Langmuir*, **2012**, *28*, 12209–12215.
85. Sandberg, M.; Eriksson, L.; Jonsson, J.; Sjöström, M.; Wold, S. New Chemical Descriptors Relevant for the Design of Biologically Active Peptides. A Multivariate Characterization of 87 Amino Acids. *J. Med. Chem.* **1998**, *41*, 2481-91.
86. Dobchev, D.A.; Mager, I.; Tulp, I.; Karelson, G.; Tamm, T.; Tamm, K.; Janes, J.; Langel, U.; Karelson, M. Prediction of Cell-Penetrating Peptides using Artificial Neural Networks. *Curr. Comput. Aided Drug. Des.* **2010**, *6*, 79–89.
87. Sanders, W.S.; Ian Johnston, C.; Bridges, S.M.; Burgess, S.C.; Willeford, K.O. Prediction of Cell Penetrating Peptides by Support Vector Machines. *PLoS Comput Biol.* **2011**, *7*, e1002101.
88. Gautam, A.; Chaudhary, K.; Kumar, R.; Sharma, A.; Kapoor, P.; Tyagi, A. *In silico* Approaches for Designing Highly Effective Cell Penetrating Peptides. *J. Transl Med.* **2013**, *11*, 74.
89. Holton, T.A.; Pollastri, G.; Shields, D.C.; Mooney, C. CPPpred: Prediction of Cell Penetrating Peptides. *Bioinformatics.* **2013**, *29*, 3094–3096.
90. Chen, L.; Chu, C.; Huang, T.; Kong, X., Cai, Y.-D. Prediction and Analysis of Cell-Penetrating Peptides Using PseudoAmino Acid Composition and Random Forest Models. *Amino Acids.* **2015**, *47*, 1485-93.
91. Wei, L.; Tang, J.; Zou, Q. SkipCPP-Pred: An Improved and Promising Sequence-Based Predictor for Predicting Cell-Penetrating Peptides, *BMC Genomics*, **2018**, *18*(Suppl 7), 742.
92. Wei, H.H.; Yang, W.; Tang, H.; Lin, H. The Development of Machine Learning Methods in Cell-Penetrating Peptides Identification: A Brief Review. *Curr Drug Metab.* **2019**, *20*, 217-223.
93. Pandey, P.; Patel, V.; George, N.V.; Mallajosyula, S.S. KELM-CPPpred: Kernel Extreme Learning Machine Based Prediction Model for Cell-Penetrating Peptides. *J. Proteome Res.* **2018**, *17*, 3214-3222.
94. Su, R.; Hu, J.; Zou, Q.; Manavalan, B.; Wei, L. Empirical Comparison and Analysis of Web-Based Cell-Penetrating Peptide Prediction Tools. *Brief. Bioinform.* 2019, 60 doi: 10.1093/bib/bby124.
95. Martelli, P.L.; Savojardo, C.; Fariselli, P.; Tasco, G.; Casado, R. Computer-Based Prediction of Mitochondria-Targeting Peptides. *Methods Mol Biol.*, **2015**, *1264*, 305-20.

96. Rondon-Villarreal, P.; Sierra, D.A.; Torres, R. Machine Learning in the Rational Design of Antimicrobial Peptides. *Curr Comput Aided Drug Des.*, **2014**, *10*, 183-90.
97. Lee, M.W.; Lee, E.Y.; Ferguson, A.L.; Wong, G.C.L. Machine Learning Antimicrobial Peptide Sequences: Some Surprising Variations on the Theme of Amphiphilic Assembly. *Curr Opin Colloid Interface Sci.*, **2018**, *38*, 204-213.
98. Mousavizadegan, M.; Mohabatkar, H. An Evaluation on Different Machine Learning Algorithms for Classification and Prediction of Antifungal Peptides. *Med Chem.*, **2016**, *12*, 795-800.
99. Grisoni, F.; Neuhaus, C.S.; Hishinuma, M.; Gabernet, G.; Hiss, J.A.; Kotera, M.; Schneider, G. *De novo* Design of Anticancer Peptides by Ensemble Artificial Neural Networks. *J. Mol. Model.* **2019**, *25*, 112 .
100. Shoombuatong, W.; Schaduangrat, N.; Pratiwi, R.; Nantasenamat, C. THPeP: A Machine Learning-Based Approach for Predicting Tumor Homing Peptides. *Comput Biol Chem.*, **2019**, *80*, 441-451.
101. Andreu, D.; Torrent, M. Prediction of Bioactive Peptides Using Artificial Neural Networks. *Methods Mol Biol.*, **2015**, *1260*, 101-18.
102. Li, Z.; Yang, Y.; Faraggi, E.; Zhan, J.; Zhou, Y. Direct Prediction of Profiles of Sequences Compatible with a Protein Structure by Neural Networks with Fragment-Based Local and Energy-Based Nonlocal Profiles. *Proteins*, **2014**, *82*, 2565-73.
103. Koch, M. H.; Vachette, P.; Svergun, D.I. Small-Angle Scattering: A View on the Properties, Structures and Structural Changes of Biological Macromolecules in Solution. *Q. Rev. Biophys.* **2003**, *36*, 147-227.
104. Petoukhov, M. V.; Svergun, D. I. Global Rigid Body Modeling of Macromolecular Complexes against Small-Angle Scattering Data. *Biophys. J.* **2005**, *89*, 1237-1250.
105. Bernadó, P.; Mylonas, E.; Petoukhov, M.V.; Blackledge, M.; Svergun, D.I. Structural Characterization of Flexible Proteins using Small-Angle X-ray Scattering. *J. Am. Chem. Soc.* **2007**, *129*, 5656-5664.
106. Svergun, D.I.; Petoukhov, M.V.; Koch, M.H.J. Determination of Domain Structure of Proteins from X-ray Solution Scattering. *Biophys. J.* **2001**, *80*, 2946-2953.
107. Mertens, H.D.T.; Svergun, D.I. Structural Characterization of Proteins and Complexes Using Small-Angle X-ray Solution Scattering. *J. Struct. Biol.* **2010**, *172*, 128-141.
108. Bizien, T.; Durand, D.; Roblina, P.; Thureau, A.; Vachette, P.; Pérez, J. A Brief Survey of State-of-the-Art BioSAXS. *Protein Pept. Lett.* **2016**, *23*, 217-231.
109. Pedersen, J.S. Analysis of Small-Angle Scattering Data from Colloids and Polymer Solutions: Modeling and Least-Squares Fitting. *Adv. Colloid Interface Sci.*, **1997**, *70*, 171-210.
110. Li, T.; Senesi, A.J.; Lee, B. Small Angle X-ray Scattering for Nanoparticle Research. *Chem. Rev.* **2016**, *116*, 11128-11180.
111. Alves, C.; Pedersen, J.S.; Oliveira, C.L.P. Modelling of High-Symmetry Nanoscale Particles by Small-Angle Scattering. *J. Appl. Crystal.* **2014**, *47*, 84-94.
112. Bressler, I.; Kohlbrecher, J.; Thunemann, A. F. SASfit: A Tool for Small-Angle Scattering Data Analysis using A Library of Analytical Expressions. *J. Appl. Crystallogr.* **2015**, *48*, 1587-1598.
113. Cuylen, S.; Blaukopf, C.; Politi, A.Z.; Müller-Reichert, T.; Neumann, B.; Poser, I.; Ellenberg, J.; Hyman, A.A.; Gerlich, D.W. Ki-67 Acts as a Biological Surfactant to Disperse Mitotic Chromosomes. *Nature*, **2016**, *535*, 308-12.

114. Sun, X.; Kaufman, P.D. Ki-67: More than a Proliferation Marker. *Chromosoma*. **2018**, *127*, 175–186.
115. Scholzen, T., Gerdes, J. The Ki-67 Protein: from the Known and the Unknown. *J Cell Physiol*. **2000**, *182*, 311-22.
116. Kozma, D.; Simon, I.; Tusnády, G.E. PDBTM: Protein Data Bank of Transmembrane Proteins after 8 Years. *Nucleic Acids Res*. **2013**, *41*(Database issue), D524-9.
117. Rice, P.; Longden, I.; Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet*. **2000**, *16*, 276-7.

TOC Graphic

