

A new biology of cell penetrating peptides

John Howl, Sarah Jones

Research Institute in Healthcare Science, University of Wolverhampton, Wulfruna Street,
Wolverhampton, WV1 1LY, UK

Correspondence

J. Howl: Research Institute in Healthcare Science, University of Wolverhampton, Wulfruna
Street, Wolverhampton, WV1 1LY, UK.

Email: J.Howl@wlv.ac.uk

Abstract

As pharmacokinetic modifiers, cell penetrating peptides (CPPs) have proven utility for the delivery of otherwise impermeable cargoes into the discrete intracellular compartments of eukaryotic cells. Methods for the molecular optimisation of CPP sequences can significantly increase the performance of novel vectors matched to a specific delivery function. Moreover, the “information rich” composition and inherent molecular flexibility of many CPPs can also facilitate their interaction with intracellular proteins and other cellular structures. This proteomimetic property of CPPs has been exploited in the design of bioportides, bioactive CPPs that regulate cellular functions often by binding relatively flat PPI interfaces to achieve a dominant negative action. This new biology of CPPs is rapidly gathering momentum through the design and synthesis of a bewildering variety of peptides that may be, in whole or part, linear, helical, cyclic and/or chimeric in nature. A particular emphasis of contemporary CPP-centred drug discovery is the unmet medical need of cancer, though the biomedical scope of bioportide applications is impressively broad including cellular signalling and reprogramming. We are hopeful that one or more of the fascinating studies described herein will translate into the clinic to establish bioportide technologies as a viable option for drug discovery.

KEYWORDS

bioportide, drug discovery, cancer, apoptosis, cellular reprogramming.

1 INTRODUCTION

*These are the words of expectation
These are the words of success, expectation
The whole world is large, and whatever it is
The whole trend is effective*

Mark E. Smith, *The Fall, Words of Expectation*, c. 1983.

Cell penetrating peptides (CPPs), mostly cationic sequences of 8-18 amino acids, are a remarkable category of biomolecule with proven utility for the intracellular delivery of otherwise impermeable cargoes. Indeed, and as reviewed,^[1,2] the therapeutic potential of CPP vectors as pharmacokinetic modifiers is widely recognised. One particular emphasis of CPPs for therapeutic application is human cancer, whence strategies can include the covalent complexation of a CPP and a tumour targeting peptide to improve the cell- and tissue-specific distribution of anticancer drugs such as doxorubicin and paclitaxel.^[3] Alternatively, the CPPs HR9 and Pep-1 have been investigated as vectors to promote the cellular uptake of a genetically modified nitroreductase for applications in directed enzyme prodrug therapy.^[4] The efficient CPP-mediated delivery of other proteins, including Cre recombinase and CRISPR associated protein, has obvious potential for genome modification and transgenesis.^[5] Moreover, the enhanced cellular uptake of nucleic acids, including antisense oligonucleotides^[6] and small interfering Ribonucleic Acid^[7-9], is another particularly productive avenue for CPP technologies. Both direct conjugation and non-covalent complex formation have been employed to combine peptide vectors and therapeutic nucleotides. There is also a bewildering variety of nanoparticulate platforms to which the conjugation of CPPs, including octa-arginine (R8)^[10] and stearylated penetratin,^[11] has improved the pharmacokinetic parameters of intracellular uptake, oral bioavailability and penetration into 3D cancer spheroids.

Concurrent with the now numerous applications of CPP vectors, of which only a small fraction are alluded to above, there has also been an equally fascinating development of bioactive CPPs.^[12-13] In the simplest sense, a bioactive CPP is a peptide that has the inherent capacity to cross the plasma membrane, reach an intracellular site and modulate some aspect of cellular physiology. **This particular scientific arena, conveniently located at the interface of chemistry and biology, is the major focus of this review. Our objectives are to provide a succinct overview of many fascinating contemporary developments and yet provide sufficient detail to encourage others to pursue their own CPP-based research strategies.**

One can postulate that a CPP-siRNA conjugate is biologically active, but in such a construct the CPP vector sequence is reasonably expected to be entirely inert. Conversely, a bioactive CPP, perhaps one that mimics a cationic helix involved in protein-protein interactions (PPIs), can not only efficiently enter a cell but could also act by a dominant-negative mechanism to disrupt a signalling interactome (Figure 1A, D). The term bioportide has been employed to describe this sub-set of CPPs^[14-15] and this designation will, when appropriate, be continued herein. Hence, and very deliberately, this is not a review of common CPP vectors, CPP-containing nanoparticles or potential mechanisms of CPP import into cells. Instead, one focus is the identification and sources of new CPPs and the experimental approaches that can be employed to improve the bioavailability and other functional parameters of selected bioportides. Methods to assess the cellular penetration of CPPs and to determine their intracellular site of accretion are also discussed, promoting a broader appraisal of developments in their biological applications. A paradigm shift away from the common view of CPPs solely as pharmacokinetic modifiers towards this *new biology* may accelerate the discovery and development of bioportides as innovative biologics and potential therapeutic agents.

2 IDENTIFICATION AND SOURCES OF CPPs AND BIOPORTIDES

2.1 PROTEIN-DERIVED SEQUENCES

Key events leading to the discovery of CPPs, beginning in the late 1980's, are amply described in Ülo Langel's latest volume.^[16] These efforts culminated in the identification of two now widely employed CPP vectors, namely penetratin (RQIKIWFQNRRMKWKK)^[17] and Tat (GRKKRRQRRPPQ),^[18] located within the primary sequences of insect- and virally-encoded transcription factors (Figure 1A). Many variations of these sequences have proven utility for cargo delivery and it is likely that these polycationic domains enable their native proteins to traverse biological membranes and fulfil their biological roles as transcriptional activators. There was no obvious *rational* methodology employed to refine either these sequences or many others later identified within human arginine-rich proteins,^[19] rather more a process of... "if it looks like a CPP it might actually be a CPP". The now expansive CPP research arena might well have developed in a rather different direction if one or more of these archetypical protein-derived CPPs had proven to possess a desirable biological activity. A pool of 272 proteins from 133 viruses proteins are a collective source of 2444 potential CPP sequences^[20] and a search for bioportides within proteins that regulate eye development in planarians^[21] also identified several very efficient CPPs including the octadecapeptide Djeya1(*H*-RKLAFRYRRIKELYNSYR-*NH*₂). Toxins, perhaps not surprisingly considering their biological functions, are also a convenient source of novel CPPs (Figure 1A). Thus, a chimeric combination of a Latarcin 1-derived spider toxin peptide (KWRRKLKCLR) with a nuclear localisation sequence from Simian virus T antigen yielded a novel CPP LDP-NLS (KWRRKLKCLRPKKKRKY), capable of protein delivery.^[22] Various linear fragments of crotamine, a cationic antimicrobial toxin polypeptide from the

rattlesnake *C. durissus*, including C1 (KQCHKKGGHCFPKEKIC), are also effective CPP vectors.^[23] Of the many hundreds of CPPs described to date^[16] a majority are now mostly confined to the annals of history; they provide structural clues useful for the development and prediction of more efficient CPP vectors, but many are essentially evolutionary dead-ends.

2.2 CPP PREDICTION AND *IN SILICO* OPTIMISATION OF NOVEL SEQUENCES

A scientifically rational approach for the identification of potential CPP sequences may be a more desirable first step towards CPP and bioportide identification. The publically accessible CPPsite 2.0^[24] is a repository of ~1850 validated CPP sequences and a valuable source of information regarding model systems and the range of cargoes delivered by individual sequences. One valid approach toward CPP identification is to employ a computer algorithm to highlight CPP sequences within the primary sequences of whole proteins and, most especially, within sequences enabling PPIs which often contain one or more cationic helices.^[25,26] The first application of this approach was described in 2005^[27] using expanded z-scales of amino acid properties to calculate four QSAR variables. The limits of two of these variables, trained by known CPPs, are particularly useful for CPP prediction. The web sites CellPPD^[28] and SkipCPP-Pred^[29] are alternative and freely-accessible platforms which employ computational methods to predict and modify CPP sequences.

It is perhaps no surprise that machine learning, a sub category of artificial intelligence, is now being utilised to expand the possibilities for CPP prediction based upon parameters including AA composition and other physicochemical properties. As an example, Diener and co-workers^[30] developed a computational prediction method based upon the properties of individual AA s but also including calculations of biochemical parameters such as mean charge, hydrophobicity, isoelectric point, water-octanol partition and α -helical

content. Such an approach enabled the characterisation of multifunctional CPPs, some of which also possess DNA-binding and antimicrobial activities.^[30] A similar approach can be publically accessed on the webserver C2Pred,^[31] an *in silico* platform which employs machine learning to predict CPPs with a potential accuracy of 83.6%.

Alternative approaches towards the identification of optimised CPPs are based upon the constructions and evaluation of larger CPP libraries. For example, comparative data derived from the evaluation of 64 CPP-conjugated phosphorodiamidate morpholino oligonucleotides (PMOs) were used to train a random forest classifier and so select CPPs specific for PMO delivery.^[32] The evaluation of seven novel CPP sequences with improved kinetics^[32] further highlights the tremendous potential of such approaches to identify enhanced CPPs and to optimise specific functional parameters such as membrane insertion.^[33] Synthetic molecular evolution (SME), an iterative process employed to design and screen combinatorial libraries exploring the sequence space around known templates, has also been applied to hybrid CPPs.^[34] The screening of a CPP library containing 8192 Tat/penetratin hybrid peptides coupled to an 18-residue peptide nucleic acid identified gain of function daughter Tat- and penetratin-related sequences. The ability of these CPPs to translocate a range of cargoes were one to two orders of magnitude more efficient than their parent CPPs.^[34] Individually and collectively, these studies illustrate the significant advantages that *in silico* computational methods and/or SME can provide for the identification and selection of most appropriate CPP vectors. These developments also hint that the same approaches might finally enable the first clinically approved CPP formulation. We can reliably anticipate that similar methodologies will also support the positive identification and biochemical optimisation of bioportides.

3 MOLECULAR DESIGNS FOR BIOPORTIDES

Monomeric linear peptides, including a majority of known CPP sequences, are mostly excellent substrates for peptidases. Thus, whilst CPPs can readily enter monolayers of cultured eukaryotic cells within a time scale of minutes, applications *in vivo* are seriously limited by the requirement of peptides to cross multiple biological barriers, achieve an effective cellular or tissue distribution and subsequently produce a biological effect in the absence of toxicity. One excellent review by Kalafatovic and Giral^[35] summarises a range of strategies which can be employed to overcome the pharmacokinetic limitations of linear CPPs, including those described below. With more selective regard to bioportides, both sychnologic and rhegynologic synthetic strategies (Figure 1B) can be distinguished.^[14,25,26] The molecular organisation of chimeric peptides is necessarily sychnologic whence distinct “domains” can be recognised within extended linear sequences. Cyclisation, a strategy to enhance the properties of peptides that bind protein targets in a helical conformation, can be applied to any bioportide including proteomimetic polycationic sequences that are intrinsically cell penetrant and so essentially rhegnylogic in nature.^[14,25,26]

3.1 CHIMERIC PEPTIDES

A variety of chemically diverse CPPs have been employed in the construction of chimeric bioportides. Thus, peptides derived from the endoplasmic reticulum protein 44 (ERp44) binding domains of adiponectin and immunoglobulin IgM, exemplified by CPP-WT (YGRKKRRQRRRKGTCAGWMA), were synthesised with an ER-targeting Tat-related CPP (YGRKKRRQRRR) as an N-terminal extension.^[36] Additional examples of the inclusion of Tat-related CPPs in chimeric bioportides include the 27 AA Tat-K, combining a short

sequence derived from the Kinase D-interacting subunit of 220 kDa protein separated from Tat by a di-prolyl linker (YGRKKRRQRRRPPNNTTTLNLPTNVRN).^[37] TAT-NLS-BLBD-6 (Tat-NLS-ATDEMIPF)^[38] is a trimeric construct consisting of Tat, a nuclear location signal (NLS) and a short sequence derived from the β -catenin/lymphoid enhancer factor-1 binding domain. The STAT6-IP bioportide^[39] is constructed from another Tat variant sequence (YARAAARQARA) as a N-terminal extension of a short 7 AA sequence surrounding phosphotyrosine⁶⁴¹ of the murine STAT-6 transcription factor.

Recently reported penetratin-containing bioportides include CP-d/n-ATF5-S1^[40]. This remarkably large 67 AA trimeric bioportide includes penetratin at the N-terminus, a dominant negative repeat sequence mimicking the DNA binding domain of Activating Transcription Factor 5 (ATF5) and a truncated human ATF5 basic leucine zipper domain at the C-terminus. RT53^[41] is a chimeric bioportide consisting of penetratin in the N-terminal covalently coupled to 37 AA sequence spanning the heptad leucine repeat domain of the Antiapoptosis Clone 11 protein (AAC-11).

The general modular nature of peptides readily facilitates the covalent additions of homing or targeting peptides within a chimeric construct (Figure 1B). As reviewed,^[42,43] both natural ligands and novel sequences derived from combinatorial screenings can enhance the on-site delivery of both peptides and nanocarriers. Some tumour-targeting peptides, including LyP1, exhibit cell penetrating properties. Hence, there is obvious potential to include homing peptides within the design of chimeric CPPs for targeted drug delivery and to specifically direct bioportides to tissues, cells, organelles and even intracellular proteins.^[42-44]

3.2 HELICAL, STAPLED AND CYCLIC PEPTIDES

There are many cogent reasons to promote the helicity of CPPs and bioportides. A detailed evaluation of helix-stabilised CPPs revealed a positive correlation between helicity and cellular uptake.^[45] Many proteomimetic bioportides derive from helical domains of proteins, particularly the molecular interfaces of PPIs. Indeed, one or more cationic helices are a common secondary structure within PPI sites where the abundance of arginine is enriched.^[25,26] The inclusion of α,α -disubstituted amino acids within the sequence of a linear peptides is a common strategy to increase the propensity for helix formation.^[26] The substitution of Ala with α -Aminoisobutyric acid (Aib), a relatively frequent modification, forces peptides to adopt either α - or 3_{10} -helical conformations due to the steric influence of the α,α -dimethyl groups. Aib might also be substituted for other AAs particularly those with similar aliphatic side chains: - Val, Ile and Leu. Aib-rich, helix-stabilised CPPs, alternatively termed foldamers, are efficient vectors for the delivery of siRNA.^[46] Similarly, chimeric CPPs, combining Aib-substituted amphipathic helical peptides with a cyclic RGD-related homing sequence at the C-terminal, can deliver anti-luciferase siRNA into human lung carcinoma A549 cells.^[47] With particular relevance to the development of rhegnylogically-organised bioportides, peptides in which cationic residues, particularly Arg, are involved in both membrane translocation and biological activity, cationic α,α -di-substituted AAs have also been developed.^[45,48] Thus, CPPs containing helicogenic AAs that are mimetics of both lysine (Ape^{C2NH₂}) and arginine (Ape^{C2Gu}) are more efficient vectors, able to more effectively deliver plasmid DNA, than their native non-helical precursors.^[48] A similar tactic to promote the helicity of polycationic bioportides could well engender improvements in uptake efficiency, *in vivo* stability and biological efficacy (Figure 1B).

The introduction of a chemical staple to maintain a helical peptide conformation can be a useful strategy towards the development of biopeptides which interact with intracellular proteins in a helical conformation. Indeed, when compared with conventional small molecular weight drugs, peptides have an increased propensity for the binding of larger protein surfaces located within PPIs. Many different chemical modifications have been utilised to induce conformational restrictions into linear peptides in an effort to enhance cellular uptake, improve proteolytic stability and promote target binding to achieve a desirable biological outcome (Figure 1B).

Predictably, cancer is a predominant target for the development of peptide-based therapeutics based upon the highly promising stapled-peptide platform. For example, Dietrich and co-workers^[49] have reported the development of a cell permeable stapled peptide inhibitor that targets the interaction between β -catenin and T cell factor/lymphoid enhancer-binding factor transcription factors, crucial elements in Wnt signalling pathways. In this later study,^[49] the stapled StAx peptide (RRWPRXILDXHVRRVWR), containing an all-hydrocarbon staple formed by ruthenium-modified olefin metathesis of two α -methyl, α -alkenyl amino acids (X),^[50] was further modified by chimeric extension with an NLS and substitution of all six Arg residues with homoarginine. The resulting NLS-StAx-h,^[49] exhibits much improved cellular uptake, compared with StAx, and is a selective and efficient inhibitor of β -catenin/transcription factor interactions. Molecular dynamics simulations of peptide-estrogen receptor binding were utilised to optimise the biological properties of a stapled peptide inhibitor (SRC-SP4)^[51] of estrogen receptor/coactivator interactions.^[51,52] The resulting peptide, R4K1, (Ac-RRRRKXLHRXLQDS-NH₂) contains a side chain to side chain olefin crosslink between branched residues (X) designed to more closely mimic the hydrophobic AA Ile⁶⁸⁹ and Leu⁶⁹³ of the I⁶⁸⁹LXXL⁶⁹³L motif of the steroid receptor coactivator 2 (SRC2) protein essential for binding estrogen receptor α . The addition of four

Arg residues at the N-terminal of R4K1 was designed to replicate the common structural motifs of amphipathic CPPs and so improve cellular permeability.^[52]

The biologically active form of trypanothione reductase from *Leishmania infantum* (Li-TryR) is dimeric and, thus, the binding interface of Li-TyrR represents a valid target for peptide PPI inhibitors.^[53] As prototypic inhibitors of Li-Tyr, libraries of stapled analogues, both 9-mer and 13-mer, of a proteomimetic sequence of Li-Tyr spanning the “hotspot” residue E⁴³⁶ were evaluated for *in vitro* activity against *Leishmania infantum*. In this case, amide-bridged peptides were prepared by the introduction of a lactam between Glu and Lys side chains with an i,i+4 spacing to stabilise a single helical turn. Chimeric conjugation of Li-Tyr-derived peptides to both Tat and nona-arginine (R9) CPPs enabled some analogues to display potent biological activities in cell culture whilst helix-stabilised examples were more resistant to proteases than their linear progenitors.^[53] A common theme of these studies^[49,52,53] is that the identification of stapled peptides with desirable uptake kinetics and biological activities will generally require the rigorous screening of a library of related compounds. Furthermore, it is certain that neither cytosolic nor nuclear accumulation is an intrinsic property of all stapled peptides.^[54] Hence, the introduction of a chemical staple may not always readily achieve pharmacokinetic and pharmacodynamic advantages above those of linear biopeptides.

The terms stapled and cyclic are often employed to describe peptides with similar gross architectures. Herein, the term cyclic is restricted to a description of macrocyclic CPPs and biopeptides that possess no obvious N- or C-termini (Figure 1B). As expertly reviewed,^[55] a majority of cyclic peptides violate Lipinski’s Rule of 5 and are accordingly unable to permeate cellular membranes. However, the discovery of a family of small, amphipathic, highly efficient and metabolically stable cyclic CPPs, exemplified by cyclo(D-Phe-Nal-Arg-D-Arg-Arg-D-Arg-Gln), Nal = L-2-naphthylalanine, facilitated more recent

discoveries of cyclic bioportides, active at PPIs, either derived from natural sources or by rational design and the screening of combinatorial libraries. Large number bicyclic peptide libraries are also a valuable source of new bioportides, including a moderately potent but cell-permeable inhibitor of oncogenic K-Ras.^[56] Finally, it is noteworthy that a comparison^[57] of the uptake mechanisms of chemically diverse cyclic CPPs, synthesised using native chemical ligation of N- and C-termini, revealed that cyclisation generally improved uptake efficiency but revealed different structural requirements to promote either glycosaminoglycan (GAG)-dependent endocytosis or direct membrane translocation.

4 INTRACELLULAR ACCESS AND ACCRETION

Historical and contemporary efforts to understanding how CPPs and bioportides gain access to the various intracellular compartments of eukaryotic cells are excellently reviewed elsewhere ^[1,2,16,35,58]. It is now established that various energy-dependent endocytotic pathways and direct membrane translocation, again in many different guises, may contribute to the observed internalisation of CPPs. However, the relevant contribution of these quite different mechanisms is a likely consequence of CPP sequence, cargo variety and cell type. In terms of the development of bioportides to modify cell biology, the precise mechanism of internalisation may be of less significance than the final intracellular site of accretion. Confocal analyses of the intracellular fate of chemically-diverse fluorescent CPPs and bioportides^[59] indicate that peptides do not assume a significant cytosolic concentration but are instead accreted within specific sites that can include organelles such as mitochondria, endoplasmic reticulum and nuclei (Figure 1C). Moreover, it is possible to fine-tune the intracellular distribution of peptides by adding proteomimetic sequences from nuclear pore proteins^[60] or mitochondria-bound hexokinase II (HKII).^[61] This individual intracellular

distribution of peptides almost certainly results from their differential binding to proteins or association with lipid membranes. Unfortunately, the mechanisms by which CPPs are trafficked within cells are relatively poorly resolved, neither is it certain whether the movement of CPPs across the plasma membrane is a two-way process.

4.1 MECHANISTIC INSIGHTS

The interaction of positively-charged peptides with negatively-charged components of the outer surface of the plasma membrane may be a common first step to promote their internalisation. A fuller appreciation of these processes may also facilitate the computational design and optimisation^[32,33] of improved CPP vectors and bioportides. The relative abundance of negatively-charged GAGs, including heparan sulphates, on the cell surface of eukaryotic cells most likely enables a first molecular interaction with some CPPs. Indeed, it is known that the relative abundance of GAGs, rather than a specific polysaccharide type, is crucial for the translocation of amphipathic CPPs, transportan (TP) and TP10, but not R9 or Tat, across giant plasma membrane vesicles.^[62] Furthermore, P21 (KRKKKGKGLGKKRDPCLRKYK), the heparin-binding domain from heparin-binding epidermal growth factor, promotes CPP-mediated delivery of enzymes, transcription factors, antibodies, nanoparticles and nucleic acids.^[63] The improved uptake efficiency of some cyclic CPPs and bioportides may likewise be the result of a more favourable interaction with GAGs.^[57]

Additional cell-surface interactions, allowing for multiplex modes of the internalisation of arginine-rich CPPs, have been recently reviewed by the Futaki group.^[64] The same team has also identified syndecan-4 as a receptor for the clathrin-mediated endocytosis of octa-arginine (R8).^[65] Intriguingly, and certainly reflective of the general

uncertainty regarding the mechanisms of CPP import, the direct translocation of elevated concentrations of R8 into cells occurs at sites of looser lipid packing.^[66]

4.2 IMPORT INTO COMPLEX TISSUES

An overwhelming majority of studies to analyse the import of peptides into cells have utilised monolayers of cultured cells and a CPP-conjugate that includes a reporter, often fluorescein or some other fluorescent dye such as tetramethylrhodamine (TAMRA).^[16, 24, 44] In such studies, confocal microscopy is a convenient methodology to determine the intracellular site of accretion of CPPs and biopeptides and semi-quantitative fluorescent assays also facilitate comparison of translocation efficacies. A detailed evaluation of both fluorimetric and mass spectrometric methods to determine the internalisation of a double-labelled penetratin analogue provided consistent results for which detailed protocols are available.^[67]

Moreover, it is obvious that, to influence the biology and/or pathology of living organisms, biopeptides may be required to traverse several biological membranes and accumulate within tissues at an appropriate concentration (~1-10 μM) to achieve efficient internalisation in the absence of cellular toxicity.^[1-3] Towards this goal two reports have championed very different systems with which to determine the penetration of fluorescent peptides into three-dimensional cellular environments. The first of these^[68] determined the penetration of fluorescein-CPP conjugates into tumour cell spheroids and explant cultures. Intriguingly, CPPs with high uptake activity (e.g. penetratin) were sequestered in the peripheral cells of spheroid cultures whilst less active examples (R9 variants) penetrated deeper into the tissue. These and other data suggested that the comparative evaluation of CPPs in a three-dimensional model may lead to a very different choice of most appropriate vector when compared to results obtained in cell monolayers.^[68]

The planarian *Schmidtea mediterranea*, a favourite organism to study the processes of tissue regeneration and stem cell biology, is also a convenient system in which to determine the penetration of fluorescent peptides into complex tissues.^[21] Whilst whole specimens are opaque, the decapitation of adults promotes a developing mass of differentiating cells, the blastema, from which regenerates an entire new head equipped with rudimentary eyes and nervous system. This tissue, useful from day 2-5 post surgery, is transparent and highly amenable to routine confocal analysis in which TAMRA-conjugated CPP can be observed in both epithelial and deeper mesenchymal cells. Since the planarian regeneration process is absolutely dependent upon the division and differentiation of totipotent stem cells (neoblasts), the same model system represents a viable platform to identify bioportides able to influence stem cell biology.^[21]

5 BIOLOGICAL DEVELOPMENTS

As indicated by the selected examples included below, the opportunities to exploit the unique properties of CPPs and bioportides as biological tools and innovative therapeutics directed to intracellular targets are manifold. Thus, it is highly probable that the capacity of bioportides to modulate potential drug modalities, significantly intracellular PPIs generally considered difficult targets for conventional therapeutics, will exploit the intrinsic target specificity, high potency and relative safety of ‘biologics’.^[58,69] The chemical remodelling of some of these lead compounds into peptidomimetic therapeutics, a strategy commonly applied to overcome the often rapid proteolysis of simple peptides, is likewise entirely feasible.^[70] Some CPPs including pVEC (*H*-LLIILRRRIRKQAH₂), a juxtamembrane protein segment, residues 615-632 from murine vascular endothelial-cadherin,^[71] are known to cross the blood-brain barrier to reliably enter the brain parenchyma.^[72] Thus, potential bioportide applications are not limited to peripheral cells and tissues.

5.1 CANCER

As reviewed,^[12,73] there is tremendous potential to employ bioportides to target the relatively expansive protein domains that enable discrete PPIs within cancer cells. Indeed, the scope of these new technologies includes CPP-mediated drug-delivery but also the direct modulation of intracellular targets with linear peptides, cyclic macropeptides and peptidomimetics.^[12,73]

As further illustrated below, bioportides can be employed to induce the death of cancer cells by apoptosis and to modulate more specific signalling pathways, mediated by PPIs or protein-DNA interactions, intrinsic to cancer cell transformation and survival (Figure 1D).

5.1.1 APOPTOSIS

Clinically useful peptide therapeutics for human cancers include gonadotrophin-releasing hormone (GnRH) agonists, GnRH antagonists and somatostatin agonists, agents acting at cell surface receptors.^[43] There is also tremendous potential to develop other anticancer peptides that can influence intracellular PPIs.^[12,40, 43,73] It is no surprise that apoptosis, or programmed cell death, is a major therapeutic target in cancer since malignant transformations lead to cell populations which display a loss of balance between division and death.^[40,74] A major challenge is to promote apoptosis in actively dividing cancer cells whilst sparing non-proliferative cells in a quiescent G_0 phase.

Two very different approaches^[41,61] illustrate the complexity of intrinsic apoptotic pathways accessible to bioportide technologies in cancer cells. As elaborated,^[41] the antiapoptosis clone 11 protein (AAC-11) is upregulated in most cancer cells to confer a survival advantage. PPIs with several apoptosis-related proteins utilise the heptad leucine zipper repeat region of AAC-11. A chimeric CPP (RT53), spanning the heptad leucine repeat

region domain of AAC-11 (residues 363–399) with penetratin at the N-terminal, specifically induces drug-induced apoptotic cell death in a variety of cancers.^[75] Moreover, RT53 also acts to promote cancer cell death, whilst sparing non-malignant cells, with a somewhat unexpected mechanism of plasma membrane permeabilisation. *In vivo*, RT53 inhibited the growth of human melanoma xenograft models and mediated anticancer effects in a murine tumour vaccination model.^[41]

The high rate of glycolysis observed in many cancer cells may reflect the overexpression and increased activity of mitochondria-bound hexokinase II (HKII).^[76] pHK-PAS (MIASHLLAYFFTELNAGKPILFF-*NH*₂), a synthetic biopptide combining the N-terminal 15 AA of HKII with a relatively short CPP, described as a penetration-accelerating segment (GKPILFF),^[77] enters cells by a combination of both micropinocytosis and energy-independent mechanisms.^[61] This is a somewhat surprising observation considering the net charge of pHK-PAS is just 1.1 at neutral pH. However, pKH-PAS is accreted within mitochondria where it dissociates the HKII protein from mitochondrial membranes and promotes cytochrome *c*-dependent intrinsic apoptosis, an effect more pronounced in cancer cells compared with noncancerous HEK293 cells.^[61]

5.1.2 β-CATENIN AND WNT SIGNALLING

The interaction of Wnt (Wingless and INT-1) protein ligands with membrane-bound Wnt receptors regulates the expression of many proteins critically involved in cell proliferation and differentiation.^[78] As previously reported,^[49] targeting the Wnt signalling pathway is an attractive therapeutic option but requires agents able to effectively disrupt discrete PPIs. Intrinsic to the common canonical Wnt signalling cascade, β-Catenin serves as a transcriptional activator of the TCF (T cell factor)/LEF-1 (lymphoid enhancer factor 1) family of DNA binding proteins to transduce signals from the plasma membrane to the

nucleus. In an effort to control β -Catenin hyperactivity implicated in the initiation and progression of cancer, Grossman and co-authors^[50] designed a series of stapled peptides to disrupt the interaction of β -Catenin with the helical β -Catenin-binding domains of TCF and other proteins. The most active, cell permeable, stapled, Axin (Ax)-derived peptide, aStAx-35R (*acetyl*-PEG-RRWPRXILDXHVRVWR-NH₂; PEG is polyethylene glycol, X is a site of cross link by olefin metathesis) was found to inhibit the growth of Wnt-dependent cancer cells without reducing the growth of cancer cells independent of deregulated Wnt signalling.^[50] More recently,^[49] further improvements have optimised the development of NLS-StAx-h (PKKKRKV-PEG-hhWPhXILDXHVhhVWh-NH₂), a derivative of aStAx-35R with an N-terminal NLS sequence and many Arg residues substituted with homoarginine (h), commercially marketed by TOCRIS Bioscience.

Hsieh and co-workers^[38] have described a somewhat similar approach to the disruption of β -Catenin-specific PPIs, employing Tat to deliver sequences derived from the N-terminal of LEF-1, known to bind β -Catenin, into the nuclei. One such peptide TAT-NLS-BLBD-6 (YGRKKRRQRRR-RKRRK-ATDEMIPF-NH₂) directly interacted with β -Catenin to inhibit the growth, invasion, migration and colony formation of breast cancer cells and suppressed breast tumour growth in nude mice.^[38]

5.1.3 ESTROGEN RECEPTOR/COACTIVATOR INTERACTIONS

Two related publications^[51,52] report the sequential development of a cell-permeable stapled peptide inhibitor of an estrogen receptor/coactivator interaction as a potential new class of breast cancer therapeutic. As described above, first generation peptides were modified with an olefinic staple designed to replace two non-interacting residues and mimic Leu and Ile residues within a sequence derived from steroid receptor coactivator 2 (SRC2) which

interacts with estrogen receptor α (ER α).^[52] As exemplified by SRC2-SP2 (acetyl-KKXLHRXLQDS-NH₂; X is a site of cross link by olefin metathesis), this strategy generated helical peptides able to bind ER α and inhibit the interaction of a steroid receptor coactivator peptide fragment. Molecular dynamics stimulations were subsequently employed^[51] to convert this high affinity stapled peptide with poor permeability into a biopeptide able to access the nucleus of eukaryotic cells. Specifically, an extended sequence of SRC2 was modified from KEKHK to RRRRK in an effort to increase cationic charge and add a putative NLS to R4K1 (acetyl-RRRRKXLHRXLQDS-NH₂).^[51] Significantly, R4K1 represses native gene transcription mediated by ER α to inhibit the proliferation of estradiol-stimulated MCF-7 breast cancer cells.^[51]

5.1.4 INHIBITION OF p53 AGGREGATION

Several sequences within the DNA binding domain of p53 are believed to contribute to an amyloid adhesive segment and promote p53 aggregation in cell lines.^[79,80] The septapeptide LTIITLE (p53²⁵²⁻²⁵⁸) forms amyloid-like fibrils and was modified by the inclusion of aggregation inhibitors such as Arg and Lys to develop an aggregation inhibitor.^[80] The most promising sequence (LTRITILE), proven to inhibit peptide aggregation *in vitro*, was chimerically combined with the CPP nona-arginine^[81] plus a tripeptide linker from the p53 sequence (RPI) to generate ReACp53 (RRRRRRRRRILTRITILE).^[80] ReACp53 penetrates primary high-grade serous ovarian carcinoma cells to convert mutant p53 from a punctate to a soluble form resembling wild type p53. As a consequence, ReACp53 rescues p53 transcription to restore apoptosis. Moreover, ReACp53 is well tolerated *in vivo* and halts the progression of and shrinks ovarian tumours bearing aggregation-prone p53.^[80]

5.2 CELLULAR REPROGRAMMING

As reviewed elsewhere,^[82,83] the general utility of CPP vectors has enormous potential to contribute to a promising future of regeneration therapy. Transcription factors, proteins that rely on both PPIs and protein:DNA interactions to regulate gene expression, are obvious targets for bioportide technologies.^[13] More specifically, and as described above, a cell-permeable dominant negative peptide inhibitor of ATF5 (CP-d/n-ATF5-S) triggers apoptosis in cancer cells whilst sparing normal cells.^[40,84]

The production of patient-specific human induced pluripotent stem cells (hiPSCs), generated by the intracellular delivery of reprogramming factors including transcription factors, microRNAs and small molecule inhibitors, is advantageous compared with a similar application of embryonic stem cells.^[82,83] Kaitsuka and Tomizawa have reported^[85] that the CPP-mediated transduction of three transcription factors, Pdx1, NeuroD and MafA, into mouse and human pluripotent stem cells generates a pancreatic lineage. Most intriguingly, the transfection of a variety of CPPs alone, including MPG (GALFLGFLGAAGSTMGAWSQPKKKRKV)^[86] and chimeric constructs contain Tat- and penetratin-related sequences, generated embryonic body-like cell clusters from normal adult human fibroblasts.^[87] This CPP-mediated conversion of somatic cells into multipotent cells is critically dependent upon peptide accretion within the nucleus and may result, at least in part, from the CPP-mediated dissociation of histone deacetylase 1 and lysine-specific demethylase 1 from reprogramming factors to reactivate their expression. Whilst this latter study^[87] provides a novel approach towards the production and reprogramming of hiPSCs and other stem cells, it also illustrates a potential caveat to the development of CPPs and bioportides in a clinical setting. Moreover, the same intriguing finding suggests that the distinction between inert CPP vectors and bioportides is at best blurred!

5.3 NEW BIOLOGY IN NEW DIRECTIONS

We present here a succinct summary of additional publications which investigate and exploit the tremendous biomedical potential of bioportides in their various structural guises (Figure 1D). Clearly, the broad scope of potential applications of cell permeable bioactive peptides, ranging from epigenetic modulation^[88] to the blockade of Toll-like receptor (TLR) signalling,^[89] suggests a very healthy future for such technologies.

The protein DPY30 is a histone methyltransferase complex regulatory subunit of the SET1/MLL family of H3K4 methyltransferases. Moreover, DPY30, significantly upregulated in some Burkitt's lymphoma samples,^[90] enhances the methylation activity of SET1/MLL complexes by directly binding a relatively short 14 AA C-terminal segment of the ASH2L subunit.^[88] A synchologic bioportide (TAT-HA-ASH2L), combining the CPP Tat with an HA epitope tag and ASH2L⁵¹⁰⁻⁵²⁹, bound DPY30 to block its interaction with ASH2L and significantly inhibit the growth of *MLL*-rearranged leukaemia and other haematologic cancer cells.^[88] A similar synchologic strategy was employed to generate TIP1,^[89] a bioportide combining penetratin with a sequence derived from the Toll/interleukin-1 receptor (TIR) domain-containing adaptor protein (SHCRVLLI). TIP1 specifically binds the BB loop of the TLR4-TIR domain to inhibit multiple TLR signalling pathways and has therapeutic potential for the treatment of TLR-mediated autoimmune and inflammatory diseases.^[89]

The expression of Cav3.2 T-type calcium channels is regulated by a molecular interaction with the deubiquitinase USP5 which is upregulated in chronic pain conditions.^[91] To block this interaction, Garcia-Caballero and co-workers^[91] developed a synchologic bioportide (TAT-cUBP1-USP5) utilising the Tat CPP to deliver a rather lengthy 35 AA peptide mimetic of the cUBP domain of USP5 (NLWLNLTDGSILCGRRYFDGSGGNNHAVEHYRETG) responsible for binding to

Cav3.2 channels. TAT-cUBP1-USP5 disrupts the Cav3.2/USP5 interaction to attenuate mechanical and thermal hyperalgesia.^[91]

The BRACA2 tumour suppressor protein plays a major role in the repair of DNA double-strand breaks by promoting the loading of RAD51 onto single strand DNA. Moreover, a 16 AA sequence derived from the BRC4 repeat motif of BRACA2 (BRACA2¹⁵²¹⁻¹⁵³⁶; LLGFHTASGKKVKIAK) blocks the specific interaction of this tumour suppressor protein with RAD51.^[92] Chimeric combination of this sequence with non-arginine generated a bioportide which confers selective hypersensitivity to poly ADP ribose polymerase (PARP) inhibition in cancer cells whilst sparing non-cancerous cells.^[92] Finally, Dominguez-Berrocal *et al.*^[93] have reported the properties of a series of multi-functional sychnologic bioportides designed to block the interaction of TEAD and YAP, two proteins involved in the Hippo signalling pathway. As a departure from more common practices, this study employed novel poly-lysine-containing CPPs, including NLS18 (RKRKKKKKWKKWPKKKKLD), to affect the intranuclear delivery of interfering mimetic sequences. The study identified NLS18-TEAD (NLS18-RLQLVEFSAFVEPPDAVD), a bioportide which displays an apoptotic effect in tumour cell lines by targeting a PPI within the cell nucleus.^[93]

6 FUTURE PROSPECTS AND CONCLSIONS

For many years we have been privileged to be a part of the gregarious international CPP community. The important work of many of the key players in the field, people and teams we have spoken to and learnt from on many occasions at numerous symposia, are represented herein. And yet this review also indicates that CPPs and bioportides are now a relatively common technology utilised in many different laboratories. Many of the studies we have

detailed employ bioportides, linear, helical or cyclic, to address a scientific question or achieve a distinct biological influence upon eukaryotic cell systems. Moreover, the unquestionable capacity of bioportides to modulate intracellular PPIs indicates that the scope of such studies may be constrained only by the limits of human imagination.

As documented herein, scientific publications will often conclude that this new biology of CPPs has therapeutic potential. **Certainly, and as expertly reviewed [58,94,95], the exploration of new therapeutic modalities beyond conventional peptide ligands will necessarily include CPPs and bioportides. Current progress for CPPs within pre-clinical and clinical trials is difficult to precisely judge. Of six CPP-containing formulations listed in 2019^[95], the sequences of two have not been disclosed. Moreover, the relatively early failure of many CPP formulations might relate to the common capacity of cationic peptides to promote the receptor-independent degranulation of mast cells.^[96] Thus, the stark reality is that few, if any, of the remarkable achievements documented herein will progress to any medical benefit.** Despite the many potential advantages,^[69,70] neither pharmaceutical companies nor major research funding agencies are really too enamoured with peptides and other biologics. Small molecular weight drugs, acting at “greasy” extracellular targets, stubbornly remain at the forefront of modern drug discovery programmes. Thus, obtaining the support to progress a peptide lead compound into *in vivo* studies and early-stage trials remains an almost impossible task.

But there is surely some hope for the future? As described above, several studies^[32,34,93] allude to the fact that it is possible to identify novel CPP sequences with specific functions tailored towards the delivery of a defined cargo into a particular tissue, cell or organelle. Whilst we eagerly recognise the historical significance of generic CPPs, including Tat, penetratin and the transportans,^[16] and acknowledge that such sequences remain commonly utilised in the construction of sychnologic bioportides, it may be time for

the field to adopt a more strategic approach to the selection and optimisation of CPP vectors. The identification of CPPs that are one to two orders of magnitude more efficient than Tat and penetratin^[34] provides evidence that it will be possible to optimise the delivery kinetics of clinically-useful peptides targeting intracellular proteins or PPIs. These potential advantages are important because the synthesis of sychnologic sequences (Figure 1B) appears to be the more common strategy currently adopted towards the identification of bioportides. Whilst several different systems have been investigated as a platform to determine CPP import into more complex, multi-dimensional cell masses,^[21,68] this remains a rather neglected area of research. A more detailed understanding of the fate of CPPs *in vivo* would almost certainly have a dramatic influence upon both the optimised design of bioportides and their routes of administration.

As research probes, commonly exogenously applied to monolayers of cultured cells, there is seemingly no obvious requirement for linear bioactive peptide sequences to be chemically altered. However, progress towards clinical utility will likely require chemical modifications to improve both pharmacokinetic and pharmacodynamic parameters.^[25,26,55] As evidenced by many of the studies summarised herein, no consensus is available to direct such modifications which are consistently challenging in a conventional academic environment. Drug discovery, of which bioportide technologies is surely a viable component,^[58,94,95] could and should be markedly accelerated if governments, pharmaceutical companies and funding agencies were challenged to work with academics to more freely provide financial support and expertise in drug optimisation and formulation. Perhaps then one or more of the bioportide sequences alluded to herein might progress into the clinic to exploit fully this remarkable new biology.

ACKNOWLEDGEMENTS

Our efforts to develop bioportides have been financially supported by the Michael J Fox Foundation, the British Council and, most recently, the Portuguese Foundation for Science and Technology (FCT). We are particularly grateful for many productive collaborations, especially with talented groups in both Sweden and Estonia led by Ülo Langel and our current “male contraceptive team” located at the University of Aveiro, Portugal. Figure 1 was produced using Servier Medical Art: <https://smart.servier.com>.

CONFLICT OF INTEREST

The authors have no conflicts of interests to declare.

REFERENCES

- [1] A. Dinca, W-M. Chien, M.T. Chin, *Int. J. Mol. Sci.* **2016**, 17, 263.
- [2] J. Xu, A. R. Khan, M. Fu, R. Wang, J. Ji, G. Zhai, *J. Control. Release*, **2019**, 309, 106
- [3] A. Sissanayake, W.A. Denny, S. Gamage, V. Saronjini, *J. Control. Release*, **2017**, 250, 62.
- [4] S. D. Anderson, R. J. Hobbs, V. V. Gwenin, P. Ball, L. A. Bennie, J. A. Coulter, C. D. Gwenin, *J. Funct. Biomater.* **2019**, 10, 45.
- [5] G. Rádis-Baptista, I. S. Campelo, J-E. R. L. Morlighem, L. M. Melo, V. J. F. Freitas, *J. Biotechnol.* **2019**, 252, 15.
- [6] A. F. Klein, M. A. Varela, L. Arandel, A. Holland, N. Naouar, A. Arzumanov, D. Seoane, L. Revillod, G. Bassez, A. Ferry, D. Jauvin, G., Gourdon, J. Puymirat, M. J. Gait, D. Furling, M. J. Wood, *J. Clin. Invest.* **2019**, 129, 4739.
- [7] D. Schirotli, M. J. Gómara, E. Maurizi, S. D. Atkinson, L. Mairs, K. A. Christie, D. F. Cobice, C. M. McCrudden, M. A. Nesbit, I. Haro, T. Moore, *Mol. Ther. Nuclei Acids* **2019**, 17, <https://doi.org/10.1016/j.omtn.2019.07.017>.
- [8] J. M. Freire, I. R. de Figueiredo, J. Valle, A. S. Veiga, D. Andreu, F. J., Enguita, M. A. R. B. Castenho, *J. Control. Release*, **2017**, 245, 127.
- [9] E-H. Ervin, M. Pook, I. Teino, V. Kasuk, A. Trei, M. Pooga, T. Maimets, *Stem Cell Res. Ther.* **2019**, 10, 43.
- [10] L. M. P. E. van Oppen, J. Pille, C. Stuut, M. van Stevendaal, L. N. van der Vorm, J. A. M. Smeitnik, W. J. H. Koopman, P. H. G. M. Willems, J. C. M. van Hest, R. Brock, *Eur. J. Pharm. Biopharm.* **2019**, 137, 175.
- [11] S. Zhu, S. Chen, Y. Gao, F. Guo, F. Li, B. Xie, J. Zhou, H. Zhong, *Drug Deliv.* **2016**, 23, 1980.
- [12] T. Jauset, M-E. Beaulieu, *Curr. Opin. Pharmacol.* **2019**, 47, 133.

- [13] I. Inamoto, J. A. Shin, *Pept. Sci.* **2019**, 111, e24048, <https://doi.org/10.1002/pep2.24048>.
- [14] J. Howl, S. Matou-Nasri, D. C. West, M. Farquhar, J. Slaninová, C-G. Östenson, M. Zorko, P. Östlund, S. Kumar, Ü. Langel, J. McKeating, S. Jones, **2012**, *Cell. Mol. Life Sci.* **69**, 2951.
- [15] I. Kerkis, A. R. de Brandão Prieto da Silva, C. Pompeia, J. Tytgat, P. L. de Sá Junior, *Cell. Mol. Life Sci.* **2017**, **74**, 647.
- [16] Ü. Langel, CPP, Cell-Penetrating Peptides, Springer Nature Singapore Pte Ltd, 2019. <https://doi.org/10.1007/978-98113-8747-0.3>.
- [17] D. Derossi, A. H. Joliot, G. Chassaing, A. Prochiantz, *J. Biol. Chem.* **1994**, **269**, 10444.
- [18] E. Vivès, P. Brodin, B. Lebleu, *J. Biol. Chem.* **1997**, **272**, 16010.
- [19] S. Futaki, T. Suzuki, W. Ohashi, T. Yagami, S. Tanaka, K. Ueda, Y. Sugiura, *J. Biol. Chem.* **2001**, **276**, 5836.
- [20] J. M. Freire, S. A. Dias, L. Flores, A. S. Veiga, M. A. R. B. Castanho, *Bioinformatics*, **2015**, **31**, 2252.
- [21] S. Jones, S. Osman, J. Howl, *Chem. Biol. Drug Des.* **2019**, **93**, 1036.
- [22] N. Ponnappan, A. Chugh, *Eur. J. Pharm. Biopharm.* **2017**, **114**, 145.
- [23] C. Dal Mas, D. A. Pinheiro, J. D. Campeiro, B. Mattei, V. Oliveira, E. B. Oliveira, A. Miranda, K. R. Perez, Ma. A. F. Hayashi, *Biochim. Biophys. Acta Biomebr.* **2017**, **1859**, 2340.
- [24] P. Argrawi, S. Bhalla, S. S. Usmani, S. Singh, K. Chaudhary, G. P. S. Raghava, A. Gautman, *Nucleic Acids Res.* **2016**, **44**, D1098. <https://doi.org/10.1093/nar/gkv1266>.
- [25] M. Lukanowska, J. Howl, S. Jones, *Biotechnol. J.* **2013** **8**, 918.
- [26] J. Howl, S. Jones, *Expert Rev. Mol. Med.* **2015**, **17**, e1. <https://doi.org/10.1017/erm.2014.24>.

- [27] M. Hällbrink, K. Kilk, A. Elmquist, P. Lundberg, M. Lindgren, Y. Jiang, M. Pooga, U. Soomets, Ü Langel, *Int. J. Pept. Res. Ther.* **2005**, 11, 249.
- [28] A. gautman, K. Chaudhary, R. Kumar, A. Sharma, P. Kapoor, A. Tyagi, Open source drug discovery consortium, G. P. S. Raghava, *J. Transl. Med.* **2013**, 11, 74.
- [29] L. Wei, J. Tang, Q. Zou, *BMC Genom.* **2018**, 18(Suppl. 7), 742.
- [30] C. Diener, G. G. R. Martinez, D. M. Blas, D. A. Castillo González, G. Corzo, S. Castro-Obregon, G Del Rio, *PLoS Comput. Biol.* **2016**,
<https://doi.org/10.1371/journal.pcbi.1004786>.
- [31] H. Tang, Z-D Su, H-H Wei, W. Chen, H. Lin, *Biochem. Biophys. Res. Commun.* **2016**, 477, 150.
- [32] J. M. Wolfe, C. M. Fadzen, Z-N Choo, R. L. Holden, M. Yao, G. J. Hanson, B. L. Pentelute, *ACS Cen. Sci.* **2018**, 4, 512.
- [33] S. A. Damiani, A. L. Alaofi, P. Dhar, N. A. Alhakamy, *Int. J. Pharm.* **2019**, 567, 118453.
- [34] W. B. Kauffman, S. Guha, W. C. Wimley, *Nat. Commun.* **2018**, 9, 2568.
- [35] D. Kalafatovic, E. Giralt, *Molecules*, **2017**, 22, 1929.
- [36] L. Hampe, C. Xu, P. W. R. Harris, J. Chen, M. Liu, M. Middleditch, M. Radjainia, Y. Wang, A. K. Mitra, *Br. J. Pharmacol.* **2017**, 174, 4478.
- [37] A. Gamir-Morrallia, C. López-Menéndez, S. Ayuso-Dolado, G. S. Tejeda, J. Montaner, A. Rosell, T. Iglesias, M. Díaz-Guerra, *Cell Death Dis.* **2015**, 6, e1939.
<https://doi.org/10.1038/cddis.2015.307>.
- [38] T-H. Hsieh, C-Y. Hsu, C-F. Tsai, C-C. Chiu, S-S. Liang, T-N. Wang, P-L. Kuo, C-Y. Long, E-M. Tsai, *Sci. Rep.* **2016**, 6, 19156.
- [39] B. T. Srinivasa, K. H. Restori, J. Shan, L. Cyr, L. Xing, S. Lee, B. J. Ward, E. D. Fixman, *J. Lekoc. Biol.* **2017**, 101, 519.

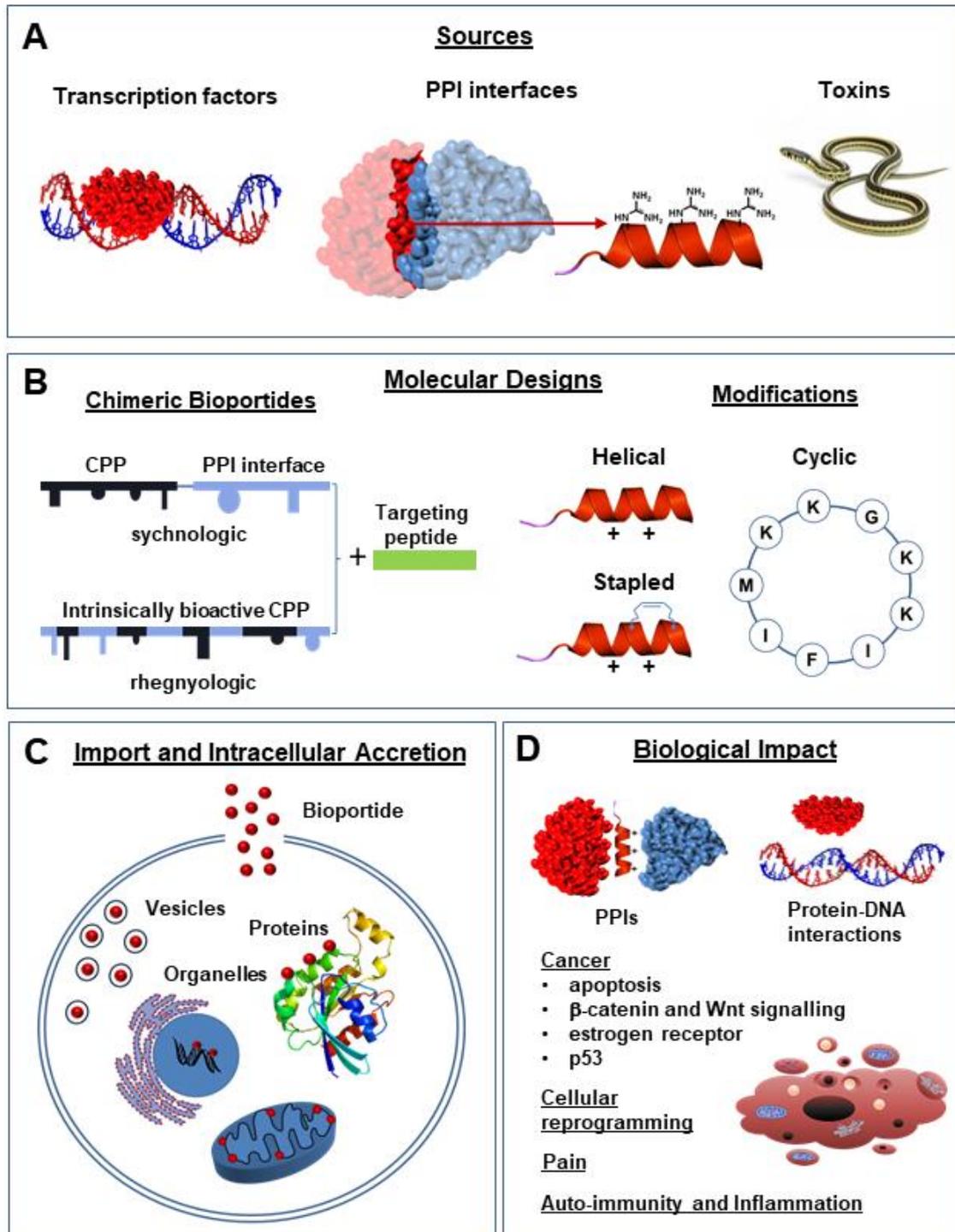
- [40] G. Karpel-Massler, B. A. Horst, C. Shu, L. Chau, T. Tsujiuchi, J. N. Bruce, P. Canoll, L. A. Greene, J. M. Angelastro, M. D. Siegelin, *Clin. Cancer Res.* **2016**, 22, 4698.
- [41] L. Jagot-Lacoussiere, E. Kotula, B. O. Villoutreix, H. Bruzzoni-Giovanelli, J-L. Poyet, *Cancer Res.* **2016**, 76, 5479.
- [42] V. Le Joncour, P. Laakkonen, *Bioorg. Med. Chem.* **2018**, 26, 2797.
- [43] K. Kurrikoff, D. Aphkhazava, Ü. Langel, *Curr. Opin. Pharmacol.* **2019**, 47, 27-32.
- [44] S. Jones, J. Uusna, Ü. Langel, J. Howl, *Bioconjugate Chem.* **2015**, 27, 121.
- [45] H. Yamashita, T. Misawa, M. Oba, M. Tanaka, M. Naito, M. Kurihara, Y. Demizu, *Bioorg. Med. Chem.* **2017**, 25, 1846.
- [46] T. Misawa, N. Ohoka, M. Oba, H. Yamashita, M. Tanaka, M. Naito, Y. Demizu, *Chem. Commun.* **2019**, 55, 7792.
- [47] S. Wada, A. Takesada, Y. Nagamura, E. Sogabe, R. Ohki, J. Hayashi, H. Urata, *Bioorg. Med. Chem. Lett.* **2017**, 27, 5378.
- [48] H. Yamashita, M. Oba, T. Misawa, M. Tanaka, T. Hattori, M. Naito, M. Kurihara, Y. Demizu, *ChemBioChem*, **2016**, 17, 137.
- [49] L. Dietrich, B. Rathmer, K. Ewan, T. Bange, S. Heinrichs, T. C. Dale, D. Schade, T. N. Grossman, *Cell Chem. Biol.* **2017**, 24, 958.
- [50] T. N. Grossman, J. T.-H. Yeh, B. R. Bowman, Q. Chu, R. E. Moellering, G. L. Verdine. *Proc. Natl. Acad. Sc. USA*, **2012**, 109, 17942.
- [51] T. E. Speltz, S. W. Fanning, C. G. Mayne, C. Fowler, E. Tajkhorshid, G. L. Greene, T. W. Moore, *Angew. Chem. Int. Ed.* **2016**, 55, 4252.
- [52] T. E. Speltz, J. M. Danes, J. D. Stender, J. M. Frasor, T. W. Moore, *ACS Chem. Biol.* **2018**, 13, 676.

- [53] M. Ruiz-Santaquiteria, P. A. Sánchez-Murcia, M. A. Toro, H. de Lucio, K. J. Guitérrez, S. de Castro, F. A. C. Carneiro, F. Gago, A. Jiménez-Ruiz, M-J. Camarasa, S. Velázquez, *Eur. J. Med. Chem.* **2017**, 135, 49.
- [54] R. Wallbrechter, P. Chéne, S. Ruetz, T. Stachyra, T. Vorherr, R. Brock, *Br. J. Pharmacol.* **2017**, 174, 2613.
- [55] Z. Qian, P. G. Dougherty, D. Pei, *Curr. Opin. Chem. Biol.* **2017**, 33, 80.
- [56] T. B. Trinh, P. Upadhaya, Z. Qian, D. Pei, *ACS Comb. Sci.* **2016**, 18, 75.
- [57] M. Amoura, F. Illien, A. Joliot, K. Guitot, J. Offer, S. Sangan, F. Burlina, *Chem. Commun.* **2019**, 55, 4566.
- [58] G. Guidotti, L. Brambilla, D. Rossi, *Trends Pharmacol. Sci.* **2017**, 38, 426.
- [59] S. Jones, J. Uusna, Ü. Langel, J. Howl, *Bioconjugate Chem.* **2016**, 27, 121.
- [60] S. Jones, T. Holm, I. Mäger, Ü. Langel, J. Howl, J. (2010) *Chem. & Biol.* **2010**, 17, 735.
- [61] A. D. Woldetsadik, M. C. Vogel, W. M. Rabeh, M. Magzoub, *FASEB J.* **2017**, 31, 2168.
- [62] J. Pae, L. Liivamägi, D. Lubenets, P. Arukuusk, Ü. Langel, M. Pooga, *Biochim. Biophys. Acta*, **2016**, 1858, 1860.
- [63] J. E. Dixon, G. Osman, G. E. Morris, H. Markides, M. Rotherham, Z. Bayoussef, A. J. El Haj, C. Denning, K. M. Shakesheff, *Proc. Natl. Acad. Sci. USA*, **2016** 113, E291.
<https://doi.org/10.1073/pnas.1518634113>.
- [64] S. Futaki, I. Nakase, *Acc. Chem. Res.* **2017**, 50, 2449.
- [65] Y. Kawaguchi, T. Takeuchi, K. Kuwata, J. Chiba, Y. Hatanaka, I. Nakase, S. Futaki, *Bioconjugate Chem.* **2016**, 27, 1119.
- [66] T. Murayama, T. Masuda, S. Afonin, K. Kawano, T. Takatani-Nakase, H. Ida, Y. Takahashi, T. Fukuma, A. S. Ulrich, S. Futaki, *Angew. Chem.* **2017**, 129, 7752.
- [67] F. Illien, N. Rodriguez, M. Amoura, A. Joliot, M. Pallerla, S. Cribier, F. Burlina, S. Sagan, *Sci. Rep.* **2016**, 6, 36938. <https://doi.org/10.1038/srep36938>.

- [68] D. vand den Brand, C. Veelken, L. Massuger, R. Brock, *Biochim. Biophys. Acta Biomebr.* **2018**, 1860, 1342.
- [69] D. J. Craik, D. P. Fairlie, S. Liras, D. Price, *Chem. Biol. Drug. Des.* **2013**, 81, 136.
- [70] N. Qvit, S. J. S. Rubin, T. J. Urban, D. Mochly-Rosen, E. R. Gross, *Drug Discov. Today*, **2017**, 22, 454.
- [71] A. Elmquist, M. Lindgren, T. Bartfai, Ü. Langel, *Exp. Cell Res.* **2001**, 269, 237.
- [72] S. Stalmans, N. Bracke, E. Wynendaele, B. Gevaert, K. Peremans, C. Burvenich, I. Polis, B. De Spiegeleer, *PLoS One*, **2015**, 10, e0139652.
<https://doi.org/10.1371/journal.pone.0139652>.
- [73] L. Mabonga, A. P. Kappo, *Int. J. Pept. Res. Ther.* **2019**, <https://doi.org/10.1007/s10989-019-09831-5>.
- [74] R. S. Y. Wong, *J. Exp. Clin. Cancer Res.* **2011**, 30, 87. <https://doi.org/0.1186/1756-9966-30-87>.
- [75] P. Rigou, V. Piddubnyak, A. Faye, J-C. Rain, L. Michel, F. Clavo, J-L. Poyet, *EMBO J.* **2009**, 28, 1576.
- [76] S. P. Mathupala, Y. H. Ko, P. L. Pedersen, *Oncogene*, **2006**, 25, 4777.
- [77] K. Takayama, I. Nakase, H. Michiue, T. Takeuchi, K. Tomizawa, H. Matsui, S. Futaki, *J. Control. Release*, **2009**, 138, 128.
- [78] H. Clevers, R. Nusse, *Cell*, **2012**, 149, 1192.
- [79] L. Goldschmidt, P. K. Teng, R. Rick, D. Eisenberg, *Proc. Natl. Acad. Sci. USA*, **2010**, 107, 3487.
- [80] A. Soragni, D. M. Janzen, L. M. Johnson, A. G. Lindgren, A. T-Q. Nguyen, E. Tiourin, A. B. Soriaga, J. Lu, L. Jiang, K. F. Fault, M. Pellegrini, S. Memarzadeh, D. S. Eisenberg, *Cancer Cell*, **2016**, 29, 90.
- [81] S. M. Fuchs, R. T. Raines, *Protein Sci.* **2005**, 14, 1538.

- [82] H. Liu, F. Zeng, M. Zhang, F. Huang, J. Wang, J. Guo, C. Liu, H. Wang, *J. Control. Release*, **2016**, 226, 124.
- [83] B. J. Seo, Y. J. Hong, J. T. Do, *Int. J. Mol. Sci.* **2017**, 18, 552.
- [84] J. M. Angelastro, *Trends Cancer*, **2017**, 3, 471.
- [85] T. Kaitsuka, K. Tomizawa, *Int. J. Mol. Sci.* **2015**, 16, 26667.
- [86] M. C. Morris, P. Vidal, L. Chaloin, F. Heitz, G. Divita, *Nucleic Acids Res.* **1997**, 25, 2730.
- [87] J. Kikuchi, N. Hayashi, N. Osada, M. Sugitani, Y. Furukawa, *Biochem. Biophys. Res. Commun.* **2019**, 518, 134.
- [88] K. K. Shah, R. H. Whitaker, T. Busby, J. Hu, B. Shi, Z. Wang, C. Zang, W. J. Placzek, H. Jiang, *Exp. Cell Res.* **2019**, 382, 111485. <https://doi.org/10.1016/j.yexcr.2019.06.030>.
- [89] H-K. Kwon, M. C. Patra, H-J. Shin, X. Gui, A. Achek, S. Panneerselvam, D-J. Kim, S-J. Song, R. Hong, K. S. Kim, Y. G. Kim, F. Y. Lee, D-H. Hahm, S. H. Lee, S. Choi, *Exp. Mol. Med.* **2019**, 51, 50. <https://doi.org/10.1038/s12276-019-0244-0>.
- [90] Z. Yang, K. Shah, T. Busby, K. Giles, A. Khodadadi-Jamayran, W. Li, H. Jiang, *J. Clin. Investig.* **2018**, 128, 3605.
- [91] A. Garcia-Caballero, V. M. Gadotti, L. Chen, G. W. Zamponi, *Mol. Pain*, **2016**, 12, 1.
- [92] A. Trenner, J. Godau, A. A. Sartori, *Mol. Cancer Ther.* **2018**, 17, 1392.
- [93] L. Dominguez-Berrocal, E. Cirri, X. Zhang, L. Andrini, G. H. Martin, S. Lebel-Binay, A. Rebollo, *Sci. Rep.* **2019** 9, 4771. <https://doi.org/10.1038/s41598-019-41404-w>.
- [94] K. Fosgerau, T. Hoffmann, *Drug Discov. Today*, **2015**, 20, 122.
- [95] K. Kurrikoff, Ü. Langel, *Expert Opin. Drug Deliv.* **2019**, 16, 1183.
- [96] M. Farquhar, U. Soomets, R. L. Bates, A. Martin, Ü. Langel, J. Howl, *Chem. Biol.* **2002**, 9, 63.

Figure 1.



Legend to Figure 1.

FIGURE 1 With a focus upon bioactive CPPs, this review presents common sources of CPPs (A), strategies for their molecular design (B) and mechanisms to study CPP import and intracellular accretion (C). Finally, we detail examples of the new biology of bioactive CPPs or bioportides (D). **Figure 1A:** Sources of protein-derived cell penetrant sequences; in addition to transcription factors and toxins, a common foundation of bioactive CPPs has been the protein-protein interaction (PPI) interface, in particular those sequences which correspond to cationic helical domains. **Figure 1B:** Molecular designs; modifications to enhance both pharmacokinetic and pharmacodynamic properties of bioactive CPPs include cyclisation, hydrocarbon stapling and incorporation of helix promoters such as α -aminoisobutyric acid. Chimeric bioportides; sequences which correspond to a PPI interface can be rendered penetrable by incorporation of a “biologically inert” CPP. This sychnologic organisation comprises a chimeric construct whereby the pharmacophores for penetration and bioactivity are distinct within the peptide chain. Alternatively, a molecular architecture that confers both membrane translocation and biological activity is typified by a rhegnylogic design, whereby the pharmacophores for cellular penetration and bioactivity are discontinuously arranged within a linear sequence. Such sequences are often derived from prediction algorithms which locate probable penetrant sequences within PPIs or entire proteins. In this conceptual diagram, the pharmacophores responsible for cellular penetration (e.g. Arg and Lys) are illustrated in black, whereas those conferring bioactivity are shown in blue. The modular nature of peptide design further enables the incorporation of a homing peptide or targeting moiety so as to ensure tissue-, cellular-, organelle- or protein-specificity. **Figure 1C:** Following membrane translocation, bioactive CPPs accrete within specific intracellular sites such as mitochondria, endoplasmic reticulum, nuclei, vesicles and proteins. **Figure 1D:** With

specific emphasis on the perturbation of PPIs and protein-DNA interactions, a selection of the latest developments in signalling targets for potential human therapies and cellular reprogramming are highlighted.