

Library construction, selection and modification strategies to generate therapeutic peptide-based modulators of protein–protein interactions

In the modern age of proteomics, vast numbers of protein–protein interactions (PPIs) are being identified as causative agents in pathogenesis, and are thus attractive therapeutic targets for intervention. Although traditionally regarded unfavorably as druggable agents relative to small molecules, peptides in recent years have gained considerable attention. Their previous dismissal had been largely due to the susceptibility of unmodified peptides to the barriers and pressures exerted by the circulation, immune system, proteases, membranes and other stresses. However, recent advances in high-throughput peptide isolation techniques, as well as a huge variety of direct modification options and approaches to allow targeted delivery, mean that peptides and their mimetics can now be designed to circumvent many of these traditional barriers. As a result, an increasing number of peptide-based drugs are reaching clinical trials and patients beyond.

Peptides as therapeutic protein–protein interaction inhibitors

Modulation of **protein–protein interactions** (PPIs) is a rapidly expanding area of interest to the scientific community. It is currently thought that there are in excess of 320,000 PPIs, with approximately 39,000 already experimentally confirmed [1,2] in humans. Many of these interactions are crucially involved in pathogenesis, and are therefore attractive targets for therapeutic agents. A large proportion of PPIs also share common structural features that are important in the design of such agents [3]; large, often flat, hydrophobic surfaces with hydrophilic points to confer specificity. These surfaces are far removed from ideal small molecule binding sites (typically solvent-accessible cavernous pockets), with small molecules often being too small to interfere significantly since they form too few specific interactions for selective targeting. As a result, whole families of potential therapeutic targets are virtually intractable to small molecule targeting. As a pertinent example, the helical secondary structure of proteins participating in helix-surface and coiled coil PPIs – which feature

in transcription factor, signaling and cell cycle protein families involved in a broad range of diseases – cannot be mimicked for inhibition of these PPIs by traditional small molecules of <500 Da [4,5]. Thus for the inhibition of many PPIs, alternative therapeutic agents are required.

To meet this need, peptides are increasingly providing excellent starting templates for PPI-interfering biomolecules. Their larger size and range of geometries allows formation of more points of interaction with the target to increase specificity and reduce potentially toxic off-target interactions. In addition, amino acid side chain interactions and their energies, torsion angles, and entropic and enthalpic contributions are increasingly well understood, meaning that it is possible to improve peptide properties predictively. This greatly improves the efficiency with which novel peptides can be isolated using high-throughput **library selection** techniques from semi-rational libraries that are tailored toward favorable properties. Finally, recent development of peptide modifications can significantly improve bioavailability and systemic stability while retaining bioactivity,

Daniel Baxter^{1,2},
Christopher G Ullman¹
& Jody M Mason^{*2}

¹Isogenica Ltd, The Mansion, Chesterford Research Park,

Little Chesterford, CB10 1XL, UK

²Department of Biology & Biochemistry, University of Bath, Claverton Down, Bath, BA2 7AY, UK

*Author for correspondence:

Tel.: +44 (0)1225 386407

Fax: +44 (0)1225 386779

j.mason@bath.ac.uk

**FUTURE
SCIENCE** part of

fsg

Key terms

Protein–protein interaction: The interaction (generally noncovalent) between two identical, similar or dissimilar proteins often with a biological consequence, which can be targeted by interfering molecules in the case of disease involvement for therapeutic benefit.

Library selection: The isolation of high-affinity target-binding library members, and the efficient removal of non-/weak binders from the library, by the application of manipulable selection pressures.

Library redundancy: Encoding of peptides with the same amino acid sequence (and thus biological activity) by multiple different DNA/RNA sequences as a result of some amino acids being encoded by multiple nucleic acid codons.

Library display: The expression of library member peptides such that the peptide is able to fold and interact with a target, while being linked to its genotype to allow peptide identification.

such that peptides and their derivatives can represent therapeutically attractive molecules [6] that increasingly are reaching the clinic [7].

Expressed peptide libraries as a source of PPI inhibitors

Peptide-encoding genetic material, utilizing either full or desired subsets of amino acids at each or select residue positions, can be expressed either *in vivo* or *in vitro* to generate a library of peptides to be ‘selected’ (panned) against the target. This permits the identification of peptide ligands with affinities exceeding that of natural binding partners, with or without prior knowledge of natural target binding molecules. Furthermore, non-/weakly interacting peptides are removed rapidly during rounds of selection to enrich target binders from the library pool by as much as ≈ 1000 -fold per round [8]. Thus library selections provide a much higher throughput alternative to individual peptide characterization using biophysical techniques, or even *in silico* modelling of target-library interactions where lack of structural information and/or difficulties in accurate molecular dynamics simulations can be problematic.

Designing peptide libraries

Peptide sequence randomization generates diversity in biophysical properties displayed by encoded peptides. Traditionally, library screening efforts have focused on maximizing diversity (number of positions randomized and/or randomization extent at each position) to increase the likelihood that a target binder will be isolated [9]. However, a major difficulty is that for highly diverse libraries, the quantity of encoded material that can be practically presented to the target, in the

range of 10^7 – 10^{14} library molecules [8,10] (the equivalent of 5–11 fully randomized residue positions), leads to incomplete library sampling. This can be partially circumvented by intelligent library design that incorporates predefined amino acids that are predicted or known to be functionally viable. These smaller and more focused libraries can be sampled with greater coverage and efficiency, as screening effort is not wasted through searching non-functional space. Libraries can also be designed to avoid peptide sequences that are known or predicted to be problematic for further development toward a therapeutic agent, for example, in the case of highly immunogenic or aggregation-promoting sequences [11]. The downside to focused libraries is that selection of serendipitous peptides with previously unpredicted binding modes might be excluded [12].

The source of this problem stems from the redundancy of the genetic code (Figure 1). This is a consequence of ‘wobble’ base pairing in the third nucleotide of a codon triplet [13], and the resulting degeneracy of nucleic acid codons in coding for the translation of amino acids. Techniques that do not actively avoid this phenomenon result in **library redundancy**; libraries contain multiple copies of identical peptides, and peptides are biased toward over-represented amino acids and against those encoded by single codons.

Library construction

Choice of library construction technique should attempt to avoid the issue of codon degeneracy to maximize selection efficiency. Traditional PCR methods such as megaprimer [14] and overlap extension [15] using NNN/NNK/NNS codons therefore are not ideal as they do not avoid these issues. Traditional ‘split-pool’ phosphoramidite synthesis approaches, featuring coupling of monomer nucleotides by phosphite triester chemistry [16], can be used for the construction of libraries; however, differential reactivity of monomers leads to unavoidable codon bias and stop codon incorporation, both of which reduce functional diversity, and achieving high degrees of randomization is expensive, time consuming and technically challenging [16,17].

Recently, a number of novel techniques avoiding the problem of codon redundancy have been developed. These include modifications to PCR-based mutagenesis (such as the ‘small-intelligent design’ approach), optimization of trinucleotide phosphoramidite synthesis components, and exploitation of enzymatic ligation to perform sequential addition of single codons (‘Slonomics™’ and ‘ProxiMAX’ technologies).

‘Small-intelligent design’ [18] is an example of a modified PCR mutagenesis approach which features the careful choice of certain degenerate codons (one per amino acid) and mixing at non-equivalent ratios. This avoids

both redundancy and stop codons, while dismissing rare codons aids near-equivalent amino acid expression levels in host systems. The main advantages of PCR-based techniques such as this are their conceptual simplicity; however, as extent of library randomization and peptide length increase, the number of mutagenic primers needed exponentially increases, quickly becoming inhibitory for libraries of reasonable diversity [17].

Alternatively, mutant libraries can be directly synthesized. Developments in solution-phase, and more recently solid-phase [19], synthesis have advanced the generation of trinucleotide phosphoramidite building blocks ('trimers') encoding all 20 amino acids on single codons where mixtures of amino acids can be specifically defined at each randomized position in a growing oligonucleotide. However, the coupling rates of different trimers are still sequence dependent and reactivity factors have to be applied in order to reduce bias toward particular codons [12,20].

Other alternatives include approaches such as Slonomics [21] and ProxiMAX [17], which feature the enzymatic ligation of codons. In Slonomics, 'Anchor' (codon-bearing) molecules are ligated via compatible single strand overhangs to acceptor framework hairpin oligonucleotides (known as 'Splinkers'). This is followed by Type IIS restriction endonuclease digestion to remove the donor framework and leave the acceptor bearing the new codon. The Type IIS enzyme cleaves outside of its recognition site and thus can be used to cleave at the same position within different library sequences. A disadvantage is that the cleavage results in a 'sticky-ended' overhang, meaning that the subsequent codon needs to be anticipated, resulting in a requirement for large numbers of oligonucleotides. An alternative process, ProxiMAX randomization [17], instead features 'blunt-end' ligation of codons carried on hairpin donor frameworks to growing acceptors to avoid this issue and reduce oligonucleotide requirement. Ligation is followed by PCR amplification, and double-strand digestion by Type IIS restriction enzyme MlyI to remove the donor framework that is then ready for addition of the next codon.

These novel techniques thus provide more precise and facile control of amino acid incorporation frequencies than traditional PCR approaches. Further, through use of single codons for each amino acid these approaches maximize sampling of the functional space within the margins of the selection systems used. A summary of these approaches is given in Table 1.

Library selection systems for peptide identification

Novel peptide sequences can be isolated from large repertoires by peptide **library display** and selection.

	U		C		A		G	
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
	UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop
	UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly

Figure 1. Redundancy of the genetic code. Shown are all possible mRNA codons from which peptides are expressed, using the four nucleobases uracil, cytosine, adenine and guanine (standard single letter code). Displayed are each triplet codon and its corresponding encoded amino acid (standard three letter code). Color of boxes denotes severity of redundancy: green denotes single encoding of an amino acid, pale pink encoding by two codons, pale orange encoding by three codons, orange encoding by four codons and finally red encoding by six codons. In library construction, use of degenerate codons results in bias in expressed peptides toward amino acids encoded by multiple codons.

Peptides (phenotype) are expressed in linkage with their encoding genetic material (genotype) and are incubated with the desired target (peptide 'display'), while selective pressures are applied in enrichment cycles to preferentially 'select' peptides with desired properties [22]. Sequencing of the genotypes of peptides remaining after selection then identifies these 'hits' for further development. High-affinity ligands can be isolated by increasing the stringency of selective pressures sequentially to isolate only the tightest binders. Additionally, libraries can be pre-selected against non-desirable targets/competitors to improve efficiency of selection of target-specific peptides [23]. A few notable examples of the numerous systems in the literature are illustrated in Figure 2. Display systems have been grouped here according to whether there is a cellular requirement to display library members or whether this can occur using the cell's transcrip-

Table 1. Comparison of library construction techniques for the generation of genetic libraries of high diversity.

Attributes	NNN/NNK/NNS	Trinucleotide phosphoramidites (codon subset)	Small-intelligent design	Slonomics	ProxiMAX
Diversity (3–12 saturated codons)	45–10%	100%	100%	100%	100%
Premature truncation (3–12 saturated codons)	11–38%	0%	0%	0%	0%
Optimization of codons to limit expression bias	×	×	×	+++	+++
Limitation of amino acid encoding bias	×	++	+	+++	+++
Codon ratio control	×	+	×	+++	+++
Codon subsets	×	+++	+++	+++	+++
Cost for 12 saturated codons	+	+++	++	+	+

Oligonucleotides synthesized by trinucleotide phosphoramidite chemistry can either be used to randomize peptides with 'NNN/NNK/NNS' randomized triplet codons (used in traditional PCR approaches), or with a subset of codons.
Adapted from [17].

tion/translation machinery in a cell-free environment that maintains a 'genotype–phenotype' linkage, as opposed to grouping according to whether panning of library against target occurs *in vitro* or *in vivo* (in cells/tissues).

In vitro cell-free display systems

In vitro cell-free systems [25,26] feature noncovalent (such as binding of proteins to the nucleic acid) or covalent chemical genotype–phenotype linkages to tether the expressed peptides to DNA or RNA. Typically, the nucleic acid libraries are translated either with cell lysate extract or recombinantly expressed and purified translation components (the 'PURE' system [27]). The expressed peptide is subsequently linked to its template so that the complex can survive intact through a panning cycle.

Cell-free noncovalent linkages

Noncovalent library systems include CIS display [8] and streptavidin-biotin (i.e., 'STA [streptavidin]-biotin linkage in emulsions'; 'STABLE') display [28] for DNA libraries, and ribosome display [29] for RNA libraries. In CIS display [8], genotype–phenotype linkage is provided by bacterial plasmid replication initiation protein RepA, specifically by taking advantage of its *cis*-activity, the high-fidelity binding of RepA to an origin of replication on the same DNA molecule from which the RepA is expressed [30]. Transcribed/translated library–RepA–DNA complexes are incubated with immobilized target and non-/low affinity binders are removed by stringent washing solutions. As with other *in vitro* screening platforms, selection pressure can be applied and manipulated easily, by altering incubation temperature, washing regimes, pH, competitive pressure (such as heterologous targets or previously selected target-binders) and/or supplementing with proteases [31].

STABLE display [28] requires each library DNA template to be terminally biotinylated and the peptides expressed in fusion with streptavidin, whose extremely high affinity for biotin directs the linkage of genotype and phenotype. Elution of DNA of selected peptide–DNA complexes for sequencing can be aided by photocleavage of a 2-nitro benzyl linker inserted between DNA and biotin.

RNA noncovalent display includes ribosome display [29]. In this system, library DNA is transcribed and translated by a cell-free expression system up to the DNA 3' terminus, where the absence of a necessary stop codon to allow ribosome release results in the ribosome with nascent library peptide remaining bound to the encoding mRNA. The major limitation of ribosome display is the stability of this mRNA–ribosome–peptide complex; conditions featuring low temperatures, high magnesium concentration and ribosome halting factors such as chloramphenicol [29] increase complex stability but also limit the range of selection pressures that can be applied to libraries.

Cell-free covalent linkages

Covalent display systems include the P2A [32] and M.HaeIII [33] systems for DNA libraries, and mRNA display [34] for RNA libraries. In the P2A system [32], expression of the P2A bacteriophage replication initiation protein is followed by covalent linkage of its catalytic tyrosine to a 5' DNA phosphate created by single strand nicking. Though a seemingly robust technique, reported inappropriate linkage to a different genotype can confuse the panning outcome. In the similar M.HaeIII system [33], library members are expressed in fusion with DNA-methyl transferase M.HaeIII, which forms a covalent bond with a fluorinated GGFC recognition sequence on the library DNA molecule. As

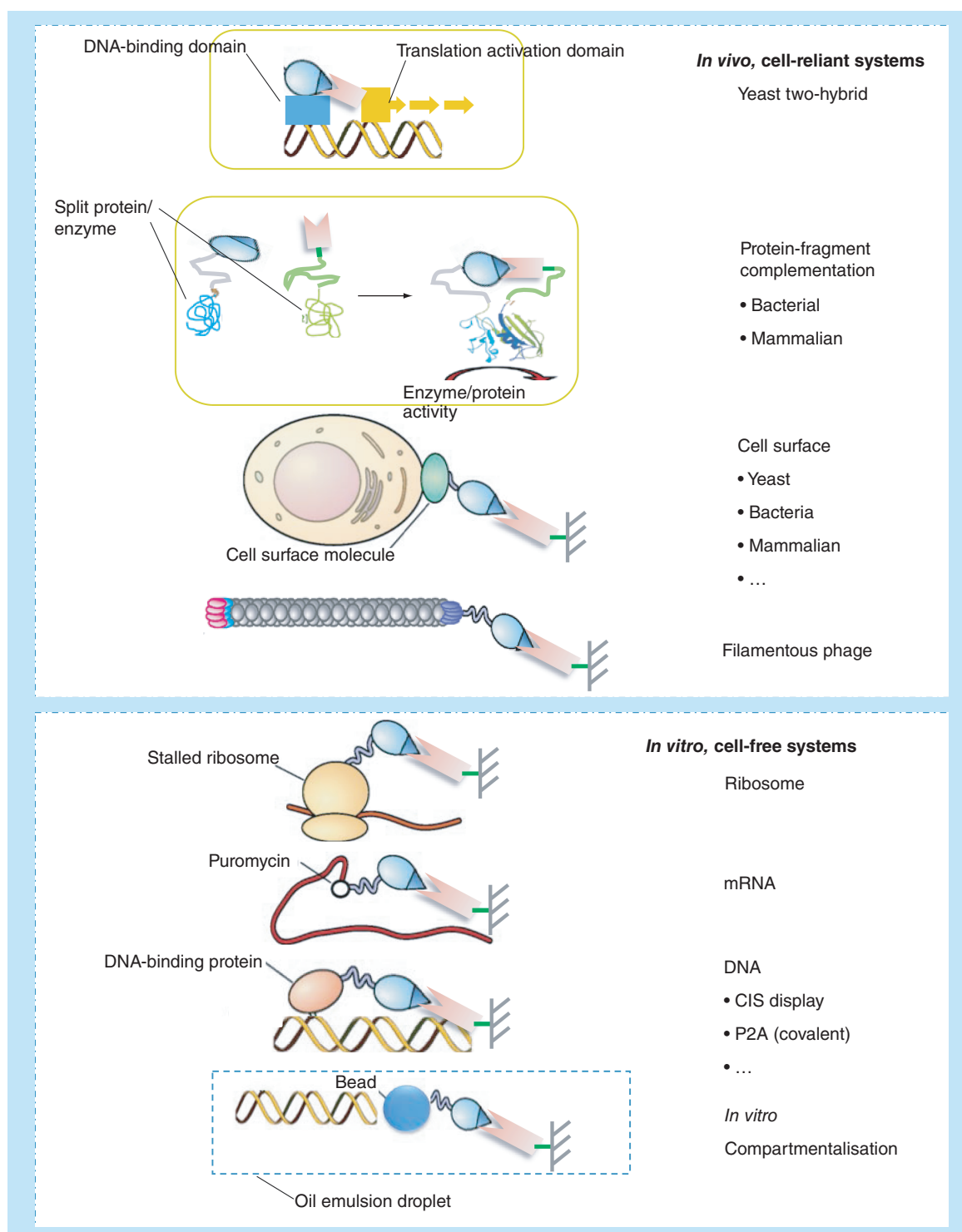


Figure 2. Library display for selection against a target in commonly used library display systems which make use of cellular expression of libraries (top panel) or *in vitro* cell-free expression (bottom panel). Target of interest, against which libraries are screened, is shown in pink. For the majority of display systems shown, the target is immobilized on a solid support (black dashed lines) via linker (green) and libraries are panned against it *in vitro*. Library-encoding DNA is shown as a yellow ribbon, or mRNA as a red ribbon, displayed peptide is shown as a light blue oval, with its target-binding site shown as a dark blue triangle. Linker molecules between displayed peptide and the molecule on which it is displayed are shown in gray. Cell walls are shown as yellow rectangles. See text for a more detailed explanation of display systems. Adapted with permission from [24] © Elsevier (2006).

for P2A, this technique suffers from free M.HaeIII peptide production, which can bind non-fluorinated recognition sequences on other DNA molecules, as well as the expense of production of fluorinated library molecules.

RNA covalent display includes mRNA display [34] as a modification of ribosome display. In mRNA display, a DNA-puromycin linker is added to mRNA library members (generated from a DNA library by transcription) by enzymatic ligation. Upon translation, the ribosome stalls at the DNA linker, allowing puromycin to enter the A site of the ribosome, whose peptidyl transferase activity moves the nascent library peptide onto the *O*-methyl tyrosine of puromycin to covalently link library mRNA to its peptide. In both ribosome and mRNA display, RNA susceptibility to degradation is a limiting factor, but library sizes achievable are much higher than techniques involving host cell transformation.

As a modification of some cell-free *in vitro* display systems, or as a stand-alone system, *in vitro* compartmentalization can be used [35], whereby femtolitre 1 μ m diameter water droplets created in an oil emulsion house a single library nucleic acid molecule which is transcribed and translated by lysate/PURE system components. Peptide–nucleic acid complexes similarly can be subjected to affinity selection in water droplets (for example displaying peptides on microbeads [36]) or as for other *in vitro* systems. These compartments help to improve the fidelity of genotype–phenotype linkages (especially noncovalent linkages, or where noncovalent interaction is followed by covalent linking) by localizing library nucleic acid molecule and expressed peptide, though compartment fusion during peptide expression is a limitation [33].

In vivo cell-dependent display systems

In vivo display systems are reliant on translation of libraries by cellular expression machinery. The majority of these systems make use of DNA library-encoding, and noncovalent linkage of genotype and phenotype. Examples include phage display [37], cell surface display [38], baculovirus display [39], protein-fragment complementation assays (PCAs) [40], yeast two-hybrid display [10] and lambda repressor displays [41]. Libraries are generated *in vitro*, individual library members transformed/transfected into individual host cells, and then peptides expressed by the cell's translational machinery. Libraries displayed on bacteriophage, cell surfaces and baculoviruses are generally panned against an immobilized target *in vitro*, while for PCAs, yeast two-hybrid and lambda repressor systems, library members that interact with a co-expressed target within the cell generate a measurable readout. For

example in PCAs, library–target interaction allows an essential enzyme/protein that is split into two fragments to refold and be activated. This may confer a competitive growth advantage [42], or generate fluorescence from re-assembly of split fluorescent proteins [43].

Phage display [37,44–45] exploits the infectivity and hijacking of host cell machinery by bacteriophage for replication, along with the packaging of bacteriophage genomic DNA within phage virions, to display peptides for affinity panning and link genotype with phenotype. In this technique, peptide display occurs in cells (*in vivo*), while affinity panning occurs *in vitro*. Phage can be split into two classes, filamentous (such as M13 phage) or lytic bacteriophages (such as lambda phage [46]), that infect a bacterial host cell and using host cell transcription/translation machinery display library peptides on phage capsids (typically also requiring provision of other necessary replication factors by helper phage infection). Capsids are either assembled in the host cell periplasmic space and secreted, or assembled in the host cell cytoplasm before cell lysis, respectively. Library peptides whose genotype is engineered into phagemid vectors are displayed as fusions with pVIII or pIII capsid coat proteins for M13 filamentous phage or gpV/gpD for lambda phage. Filamentous phage display is routinely used for small peptides, while lytic phage may be used for larger proteins whose folding is sensitive to periplasmic translocation. With both classes of phage, phage particles are then purified from bacterial cells and subjected to panning *in vitro* against an immobilized target. A major limitation to phage display relates to library sizes, and is common to all cell-reliant display systems. Libraries genetically engineered into vectors must be transformed into host cells before peptide expression, such that transformation efficiencies limit library diversities to 10^{10} – 10^{11} members. Despite this, phage display benefits from robustness of selection strategy, stability of phage virions and flexibility to adaptation, making it one of the most widely used selection strategies.

In addition to traditional phage *in vitro* panning, the extraordinary stability of phage also allows panning against whole cells from a tissue biopsy or even *in vivo* by means of intravenous injection [45,47]. This allows selection of peptides against extracellular, membrane and even intracellular targets (if peptides capable of carrying phage across cell membranes are displayed in tandem [48]) accessible to phage administration. The advantage of this technique is that peptides are selected in a real, relevant biological context that is quite different to idealized *in vitro* conditions [23]; not only is target interaction selected for under systemic pressures applied by the circulation, immune system, organs, tissues and cells, but also peptides are

selected for favorable pharmacokinetic properties such as blood retention, low immunogenicity and resistance to biological degradation.

Cell surface display technologies [38] generally comprise genetic fusion of a DNA library to the coding sequence of a protein expressed and incorporated into the cell wall, such that the fused library peptide is presented to the extracellular media. For example, in yeast surface display, library members can be fused with α -agglutinin or flocculin membrane proteins. Different cell types require different choices of cell membrane proteins, but otherwise these technologies are very similar. Once a library has been displayed on the cell surface, panning against immobilized target or other cells can be performed as for other display systems but also with the advantage of gating for high-affinity target binders using a cell sorter. These systems are advantageous where peptide library members are long and fold in a more complex way, such that natural folding can be replicated efficiently by whole-cell machineries. Other advantages are more specific to the cell type used; for example, yeast cell mating can lead to increased library diversity and library evolution by genetic recombination [49].

Baculovirus display [39] exploits Baculoviridae family invertebrate viruses to infect mammalian or other eukaryotic host cells and display peptide libraries both on host cell surfaces (thus constituting a form of cell surface display), and/or on budding virions. For budding virions, fusion proteins between library and major baculovirus envelope protein gp67 are generated by library insertion into baculoviral genomes under the strong transcriptional control of the non-essential polyhedron gene promoter. As for cell surface display, baculovirus display particularly benefits panning of libraries where peptides require eukaryotic folding/post-translational modifications.

In DHFR-PCA [40], library members are genetically fused to a rationally designed fragment of murine DHFR. Similarly, the target is fused to the complementary fragment. Library–target interaction reconstitutes DHFR which proceeds to metabolize nucleotide precursor dihydrofolate in minimally nutritious growth media to allow detectable cell growth, otherwise inhibited by the absence of available nucleotides and suppression of genomic DHFR by trimethoprim.

Yeast two-hybrid assays [10] feature splitting of the Gal4 transcriptional activator protein between its DNA binding and activation domains, and fusion of library or target to each domain. Only library–target interaction reconstitutes Gal4 to initiate transcription of a reporter gene such as a fluorescent protein, β -galactosidase for blue-white colony screening or an antibiotic-metabolizing enzyme.

Finally, in lambda repressor display [41], library peptides are expressed as fusions with the lambda repressor of transcription monomer within a host cell, with the target likewise fused to another lambda repressor monomer. Library–target interaction promotes lambda repressor dimerization, which represses transcription of *tof* immediate-early gene to inhibit lambda phage lysis of host cells such that target-interacting library members are identified by host cell survival.

In addition to selection for high affinity, display systems featuring selection in cells have the additional advantages of simultaneously selecting for target specificity, solubility in the cell environment, non-aggregation, non-toxicity, stability in a reducing environment, protease resistance and ability to outcompete endogenous competitors, both for the target and the library [40,50]. PCAs such as DHFR-PCA are attractive in that selection stringency can be increased by employing growth-restrictive cell culture conditions such that only the most effective target-binders will reconstitute enough DHFR activity for cell survival. Additionally, competitive and negative design can provide extra competition through the co-transformation and co-expression of exogenous target-binding competitors, or heterologous ‘off-targets’ [51], which selects for high target selectivity in isolated peptides.

Selection system choice

Cell-free and cell-dependent systems each have strengths and weaknesses, some of which overlap. Cell-free systems allow the expression of much larger peptide libraries (in the region of 10^{14} members [8]) relative to cell-dependent systems (a maximum of around 10^{11} members [45]) due to limitations in transformation/transfection efficiencies. Thus, cell-free systems enable greater sampling of peptide sequence space, a greater chance of identifying peptide ‘hits’ and consequently potentially more optimal hits [9]. However, false positives are a rarer occurrence in cell-reliant systems where the stringency of selection is higher [52]; cell-free systems can suffer from inadequate removal of non-specific off-target binders (often peptides that adsorb to well-plate surfaces) [53]. Sensitivity of cell-reliant systems to weaker target binders (which may nevertheless generate useful antagonists with further affinity maturation/modification) can be better than cell-free systems [54]. Finally, cell-free selection systems benefit from increased control/easier manipulation of selection stringency [31].

Ultimately, choice of selection system is largely case-dependent, and is affected by library diversity, complexity of folding/post-translational modifications

Table 2. Comparison of important attributes of some commonly used peptide display systems for use in peptide selection.												
Attribute	Cell-free expression					Reliant on cellular expression						
	Ribosome	mRNA	CIS	Covalent (P2A/M.HaeIII)	IVC	STABLE	Phage	Cell surface	Baculovirus	Yeast two hybrid	Lambda repressor	PCA
Genotype-phenotype linkage	NC	C	NC	C	NC/C	NC	Capsid protein	Membrane protein	Capsid protein	NC	NC	NC
Library nucleic acid	mRNA	mRNA	DNA	mRNA/DNA	DNA	DNA	DNA	DNA	DNA	DNA	DNA	DNA
Theoretical maximum library size	>10 ¹³	>10 ¹³	>10 ¹³	>10 ¹³	>10 ¹⁰	>10 ¹³	<10 ¹¹	<10 ¹¹	<10 ¹¹	<10 ¹¹	<10 ¹¹	<10 ¹¹
Expression of complex peptides (folding)	Limited	Limited	Limited	Limited	Limited	Limited	Limited	Eukaryotic	Eukaryotic	Eukaryotic	Eukaryotic	Eukaryotic
Target biopanning	In vitro	In vitro	In vitro	In vitro	In vitro	In vitro	In vitro	In vitro	In vitro	In vivo (cells)	In vivo (cells)	In vivo (cells)

See main text for full explanation of display techniques.
C: Covalent; IVC: *In vitro* compartmentalization; NC: Noncovalent; PCA: Protein-fragment complementation assay.
Adapted from [32].

of library peptides, as well as cost, complexity and time. Some attributes of the various display systems discussed are summarized in Table 2.

Enhancing simple library selections

Recent development of next-generation sequencing (NGS, or ‘deep sequencing’) technologies, such as Illumina MiSeq [55], Life Technologies Ion Torrent [56], single-molecule real-time sequencing [57] and Oxford Nanopore [58] technologies, has expanded sequencing capability to hundreds of thousands of reads through a library pool. This has obvious application to the quality control of constructed libraries, providing data on amino acid incorporation frequencies and frequencies of truncation/mutation, but additionally now allows the enrichment process to be followed during selections by sequencing genetic material isolated after each selection round [59]. A particularly interesting use of the data gathered during selections is to identify binding motifs common to early stage binders that later are outcompeted by other motifs. Though undesirable in this state, these motifs nevertheless may prove attractive if revisited and randomized in a future library, or may be interesting for other applications, or against homologous targets.

Library selection systems provide not just the opportunity to isolate novel interactions, but also the chance to evolve the binding characteristics of previously identified binders toward higher affinity and specificity. Directed evolution can be achieved by introducing an element of randomization at stages of selected library pool amplification between selection rounds in conventional library selection, or when constructing secondary libraries from initial hits. Randomization can be achieved using techniques [60] such as error-prone PCR, exon shuffling and reassembly via PCR, error-prone host replication (mutator strains) or recombination procedures. Thus, an initial library evolves to cover greater sequence space around binding motifs than initially possible.

Phage Assisted Continuous Evolution (PACE) [61,62] is an interesting example of modification of a display system to effect directed evolution during selection. In PACE, the M13 filamentous bacteriophage is modified such that the pIII coat protein responsible for infecting new host *E. coli* is instead placed on an accessory plasmid and transformed into host cells before selection. Expression of pIII is linked to library peptide interaction with the target, such that non-interaction blocks pIII expression and thus results in the production of non-infectious phage, which are washed away in the flow of fresh host cells supplied by a microfluidics device. Conversely, library–target interaction leads to pIII expression, allowing formation of infectious phage that infect the new host cells delivered to survive the selection round. Every successive phage replication results in mutation of the

peptides being selected as a result of host cell replication infidelities, such that they continuously evolve to arrive at sequences with the highest affinity for the target.

Directed evolution of initial hits in secondary libraries may feature randomization of arbitrary regions of a binder where a deeper understanding of binding is absent, or randomization of residue positions identified as crucial for interaction by predictive, biochemical or structural analyses. Often this means that sequence space can be more thoroughly searched than in the initial library, and consequently affinities of initially selected hits can be further matured to increase therapeutic attractiveness.

Synthetic peptide libraries as a source of PPI inhibitors

Peptide library selection strategies do not always require a genotype–phenotype linkage. Synthesized non-genetic peptide libraries make use of automated peptide synthesizers or synthesis on microarray chips [63] to synthesize hundreds of thousands of unique peptide sequences using standard 9-fluorenylmethyloxycarbonyl (Fmoc) peptide chemistry. Commercial companies, such as PEPperPRINT, make use of modified laser printers to ‘print’ amino acid building blocks as array spots before coupling to growing peptide chains [64]. Immobilized libraries are then incubated with a solution of the target molecule, whose linkage to a reporter molecule (i.e., fluorescein isothiocyanate, horseradish peroxidase etc.) allows identification of interacting peptides after removal of non-weak binders by stringent washing. Screening peptides on microarrays benefits future library design in a similar way to NGS of DNA libraries; not only are high-affinity binders identified, but also those which are similar in sequence but display lower affinities. This can provide residue-specific insights into the structure-dependence of the binding mode, from which new library designs can benefit. The advantages of peptide microarrays over genetic systems are their rapid production, and reuse for screening against multiple targets. However, currently the library complexity achievable is limited by the density of peptides that can be practically applied to a solid support (in the region of 1 million per cm² [65]), meaning genetically encoded libraries currently offer much greater complexity prior to screening.

Peptide development for therapeutic administration

Traditionally, small molecules have been considered the only tractable interfering agents for therapeutic administration as they can be highly bioavailable across biological membranes to target intracellular receptors, stable in extracellular fluids and non-immu-

nogenic [66]. However, their small size limits the number of points of interaction that can be made with protein interaction interfaces, which in turn can generate toxicity issues relating to target-specificity. Peptides, on the other hand, while large enough for high specificity, can suffer from poor stability, bioavailability, the inability to permeate biological membranes to reach intracellular targets or those beyond the blood-brain barrier for example, and reactivity with T-cell receptors and B-cell/antibody paratopes [67]. A small number of unmodified peptides that are quite large relative to small molecules and which do not adhere to ‘Lipinski’s Rule of Five’ have made it to the clinical market since the first administration of natural peptide hormones in the 1980s. However, these often still display several limitations, as exemplified by enfuvirtide [68] whose instability requires dosages on the milligram scale. As a result, the number of unmodified peptides achieving regulatory approval relative to the number entering clinical trials in this time period has been low [69]. However, in recent years the distinction between small molecules (generally <500 Da) and peptides (usually >500 Da) has narrowed with respect to their pharmacokinetic properties [66] with the continued development of existing and novel peptide modification strategies. These strategies afford conversion of peptides into ‘peptidomimetics’ which, while retaining many of the structural characteristics and high target specificity of the parental peptide, can be substantially more stable and capable of reaching even the most difficult intracellular targets, thus enabling interference of small molecule-intractable PPIs with small molecule-like pharmacokinetics. Modification strategies can be applied to peptides isolated by display and selection systems either before or after selection to provide high-throughput discovery of therapeutically attractive molecules.

Post-selection modification of peptides

For the modification of peptide hits after identification by screening strategies, a wide range of approaches have been invented and developed in the last 30 years [70,71]. Peptides can be synthesized chemically to include non-natural amino acids, or non-peptidic backbones. Additionally, ‘retro-inverso’/D-amino acid, structural constraint/cyclization, hydrogen-bond surrogacy, conjugation and truncation strategies (or combinations of the above) are available. These techniques aim to

Key term

Peptidomimetics: A small peptide-like compound with adjusted molecular properties that aim to enhance bioactivity (target-binding affinity or selectivity, or pharmacokinetic properties etc.).

improve therapeutically desirable properties such as cell penetration/bioavailability, decrease protease-susceptibility, improve half-life and increase systemic stability, to improve overall druggability.

Non-natural side-chains & backbones

Natural α -amino acids can be substituted with those bearing highly derivatized/novel side chains [72], and/or amino acid mimetics such as β - [73] or N-substituted amino acids (the latter forming 'peptoids') [74]. A wide variety of non-natural amino acids have been incorporated into peptides, with novel side chains featuring alkyne, thiazolidine, aryl halide, alkene, heterocycle, ketone, azide, metal-chelating and photochemically reactive functional groups among others [72]. One major advantage of this approach is that structural restrictions that limit the strength and complementarity of α -amino acid interactions with the target can be overcome [75]. Non-natural amino acids may also disrupt protease recognition sites (for example, β -amino acid substitution [76]) or immune cell/antibody paratopes that have evolved to recognize α -amino acids [77].

Non-peptidic backbones can replace the relatively labile and protease-scissile amide bond in peptides for a moiety more resistant to proteases, and one less recognizable to immune/antibody paratopes. Non-peptidic backbone strategies evolved at a similar time to non-natural amino acid substitution, and have included substitution with amino acid mimetic or amino acid-like building blocks (such as β -amino acid [76] or sulfonopeptide units [78]), or quite dissimilar modules more reminiscent of small molecule drugs (such as pyrrolopyrimidines [79], or aryl/imidazolidinones [80]) and may replace entire peptide backbones or only regions therein.

For α -helix mimetics, β -amino acids can be constructed into oligomers that generate conformationally similar helices, thus retaining binding ability [81], while being more resistant to protease digestion [76]. Similarly, oligoureas, oligosulfonamides and pyrrolopyrimidines are examples of moieties that are capable of adopting the necessary Φ and Ψ angles for α -helix adoption [82]. Oligoenaminones [83] and oligobenzenamides [84] also have similar mimetic properties. These moieties can display a wide variety of properties based on backbone flexibility and side chain decorations [79], in order to mimic the side chain functionalities and spatial characteristics of α -helices, with the added advantage of greater customization of side chains to offer potential optimization of binding. In addition, even partial backbone replacement can afford protease protection by disrupting recognition sites.

Partial non-peptidic backbone replacement, for example using cubic alkanes and conjoined rings, has also been used to mimic β -turns and β -sheets [85].

However, careful adaptation to ensure suitability for mimicry of turns involved in molecular recognition is necessary to ensure replication of amino acid side chain groups' geometries and interactions [85].

D- α -amino acid substitution replaces natural L-enantiomer α -amino acids with D-enantiomers, while the related 'retro-inverso' strategy reverses a fully substituted D- α -amino acid sequence from N- to C-terminus to arrive at a mirror image of the L-enantiomer [86,87]. Both techniques significantly decrease susceptibility to human proteases evolved to recognize L- α -amino acids [88], and immune system reactivity, which is primarily against L-enantiomer peptides [77]. D-peptides can be synthesized by solid-phase peptide synthesis (SPPS), aided by the continued development of more efficient methods for D-amino acid synthesis [89]. D-peptides can also be indirectly selected by library display systems such as 'mirror image phage display.' By reversing the chirality of the target molecule from L- to D-enantiomer, and selecting against a library of L-peptides, by simple logic the L-enantiomer of the target should then be bound by the D-enantiomer of the selected peptide [90]. This strategy is however only really attractive where the D-target is amenable to SPPS.

Constrained secondary structures

In the modification of peptides toward therapeutic administration, undoubtedly the two most popular strategies currently are structural constraint (particularly of helical peptides) and macrocyclization, both of which are successfully generating peptidomimetics that are reaching the clinic [7].

Structural constraints are non-natural modifications to peptide motifs that cyclise two amino acids and thus the peptide backbone [70]. In doing so, the entropic cost to adopting the particular fold of the motif [91] is reduced, such that peptides are then entropically pre-organized for binding to their target [92].

Constraint of α -helices, β -sheets/turns and loops have all been described [93,94]. The entropic pre-organization of α -helices by constraint strategies [70,95] has received particular attention, and has been shown to improve binding affinity of helical peptides of various sequences and lengths with their therapeutic targets [92,96–98]. For example, in the case of oncogenic transcription factor Activator Protein-1, whose activity is dependent on entropically unfavorable helix formation and then supercoiling of α -helical Jun and Fos proteins [99], helix constraint of a c-Fos antagonist favorably reduced the entropic penalty to antagonist–c-Fos interaction to increase binding free energy/affinity, and additionally facilitated truncation to smaller and thus more therapeutically attractive antagonists while retaining binding [92].

In another example, the hydrocarbon constraint (or

‘staple’, referring to a hydrocarbon constraint generated by olefin metathesis) of Hdm2 inhibitor SAH-p53-8, a derivative of the Hdm2-interactive p53 helix, conferred high helicity to SAH-p53-8 for effective mimicry of the p53 helix, leading to increased p53-mediated cancer cell death [100]. Interestingly, interaction of the staple itself with Hdm2 residues conferred a serendipitous benefit to binding affinity [101].

In addition to beneficial effects on target binding, helix constraint has been reported to improve membrane penetration [102,103] beyond the natural tendency for amphipathic α -helices to associate with membranes. Helix constraining also promotes and shields a protease-resistant helical conformation [104], but has the potential disadvantage that increasing peptide molecular weight may be detrimental to bioavailability.

Structural constraints can be introduced either during SPPS [105], or post-synthesis [106]. The range of constraint modalities available has grown vastly in the last 30 years to include lactam/peptide/amide bridges [92], aryl halides [106,107], hydrocarbon chains (generated by ‘ring closing metathesis’) [108] and metal chelation to force helicity [109]. These can be inserted at suitably spaced positions along the helix, with ($i \rightarrow i + 3$), ($i \rightarrow i + 4$), ($i \rightarrow i + 7$) and ($i \rightarrow i + 14$) being spatially close or planar to each other, and are usually inserted away from the interaction interface(s) on constrained peptides to avoid disruption/alteration of binding. Despite the wide variety of constraints available, and the obvious efficacy of conferring helicity on peptides as short as five amino acids [98], choice of which constraint to apply to a given sequence for a desired gain – for example, maximal α -helicity – has not previously been clear. However, a recent strictly comparative investigation goes some way to address this, indicating that lactam constraints are the most helix-inducing constraint in alanine pentapeptides, outperforming popular hydrocarbon and thioether linkages [95].

In the last year and a half, the first stapled peptide has reached clinical trials. Aileron Therapeutics’ ALRN-5281 is a stapled analog of human growth hormone-releasing hormone designed to increase natural growth hormone release for the treatment of rare endocrine disorders, and has successfully completed Phase I safety trials in healthy humans and is expected to enter Phase II trials imminently [110]. This success demonstrates the potential for helix constraint strategies to deliver therapeutically viable peptidomimetics.

Two related strategies to helix constraint are hydrogen bond surrogacy (HBS) and helix templating. Hydrogen bond surrogacy [111] involves the use of covalent hydrocarbon linkers to replace the terminal backbone-to-backbone hydrogen bonds that stabilize an α -helix, again reducing the entropic penalty to

α -helix formation [112]. A benefit over constraining is that there is no interference with amino acid side chains, which may be crucial for target interaction [111]. Helix templating [113] involves the use of synthetic terminal amino acid analogs, complexation of metal ions or multidentate rings [114], designed such that ψ and ϕ angles are fixed at helix-promoting angles, to provide a nucleus from which an α -helix can preferentially propagate.

Similar structural constraints can reinforce β -turns/strands/sheets [115], loops and reverse turns in larger peptides or even small proteins that represent ‘scaffolds’ upon which library peptides can be presented for target binding [116]. Hydrocarbons [117], synthetic small molecules, D-amino acids, non-natural amino acid mimetics, disulfides or even other proteins [93] have all been used for constraint, which further demonstrates the flexibility of constraining to suiting the structural needs of a large repertoire of specific cases. Similarly, templating can structurally reinforce β -turns/sheets and loops [114], though this is complicated by turn/sheet formation dependency on stabilizing interactions from surrounding turns/sheets, and the tendency for formation of high-order aggregates. Insertion of synthetic molecules such as epindolidione between stretches of amino acids can provide hydrogen bonds to backbone peptide groups in order to allow a β -strand to adopt a sheet conformation [118]. Alternatively, molecules that replicate the geometry of β -turns – including pyridine analogs, in some cases in coordination of metal cations – are able to reverse the direction of β -strands and bring their backbones into range for hydrogen bonds to stabilize a β -sheet conformation [114,119]. Finally, adoption of omega loops – hydrogen-bonded loops of irregular backbone dihedral angles that participate in molecular recognition, protein folding or stabilization [120] – has been achieved using naphthalene amino acid derivatives [121]. Constraint of non-helical proteins conveys benefits such as protease resistance [122] in a similar way to α -helices.

Peptide macrocycles – featuring the cyclization of one or both peptide termini either with each other or with intervening amino acid side chains – were first described in natural peptides 60 years ago, and since then much interest has been taken in their desirable pharmacokinetic properties [123]. In the last 15 years in particular, like structural constraint, macrocyclization has been explored thoroughly for peptide development [124] and has proven to be an effective strategy for improving both pharmacokinetic and pharmacodynamic properties of peptides. For example, cyclization ensures that peptides avoid linear protease-susceptible conformations [122]. Further, cyclizations may improve cell penetration by encouraging multiple intramolecu-

lar hydrogen bonds which improve the otherwise energetically unfavorable desolvation of the peptide bond necessary for membrane association [125]. Cyclization is often coupled with N-methylation for enhanced cell penetration, where N-methylation masks solvent-facing backbone polar groups requiring desolvation [126]. Cyclization can also have positive effects on pharmacodynamics of peptides where target interaction involves a strained conformation, such as improving affinity by reducing entropic loss upon target-binding, just as for peptide constraint.

As a result of their favorable characteristics, some macrocyclic peptides have already reached the clinic, while others look set to be approved for administration in the near future. For example, pasireotide [127] is a cyclic peptidomimetic of somatotrophin release inhibiting factor which has been approved for the therapy of chronic hypercortisolism [128]. Oritavancin is a promising Phase III candidate peptidomimetic analog of vancomycin, which aims to afford desperately needed treatment of vancomycin-resistant MRSA [129].

Chemical synthesis of macrocyclic peptides has been described using a number of different chemistries [124], while creative strategies are available to overcome adverse conformational restraints to macrocyclization, including the entropic penalty to folding. Chemistries include metal-ion catalyzed cyclizations, sulfur cyclizations, ring-contraction, ring-closing metathesis, 'click' chemistry [130], multicomponent reactions for direct synthesis of cyclic peptides, electrostatically controlled reactions (use of polar organic solvents to pre-organize peptides into cyclization-ready conformers) and native chemical ligation (NCL)/expressed protein ligation (EPL) [131,132], to generate macrocycles during or post-SPPS featuring thioether, disulfide, hydrocarbon, lactam and other linkages [124]. In particular, NCL/EPL is a simple technique for cyclization that generates a peptide bond, while 'click' chemistry triazolyl linkages have received much attention for their presence in natural cyclic products of interesting biological activity and stability, and ability to mimic *cis*- and *trans*-peptide bonds depending on substituent groups. NCL/EPL for cyclization proceeds by nucleophilic attack of a C-terminal cysteine sulfur on an N-terminal α -thioester which through a five-membered ring intermediate rearranges to release the cysteine side chain and form a peptide bond between N- and C-terminal amino acids [131]. Use of orthogonal protecting groups on multiple cysteine residue sulfurs of a single synthetic scaffold molecule can also allow directed sequential cyclization of multiple cysteines in a peptide by NCL [133]. 'Click' chemistry typically involves Cu(I)- or Ru(II)-catalyzed cycloaddition between an N-terminal azido-amino acid and a C-terminal

alkyne amino acid mimetic to form a 1,4-disubstituted 1,2,3-triazole link [130]. Once macrocyclic, peptides are still amenable to further modification, including sequence diversification through exploitation of aziridine amides to allow site-specific incorporation of 'foreign' peptide/peptidomimetic sequences [134].

Formulation

Peptides can be conjugated to specialized protein transduction domains [135] which include basic cell-penetrating peptides, α -helical peptides and viral fusion proteins [136], that are capable of crossing biological membranes by various mechanisms [137]. Conjugation to carrier molecules such as polyethylene glycol or the novel peptoid N-methoxyethyl glycine, or human serum proteins such as human serum albumin, slows renal clearance and thus improves tissue uptake, improves proteolytic stability and lowers immunogenicity [138,139]. Conjugation of peptides to proteinogenic moieties can be achieved simply by expression of genetically fused species, while conjugation to non-natural molecules can be achieved either post synthesis using activated molecules reactive with amino acid side chains [140], or using synthesized amino acid derivatives during SPPS [141].

Truncation can be utilized to improve cell penetration potential [50,142] to counter limitations to the available space between cell membrane proteins through which peptides can permeate. Truncation can also be targeted to remove known immune cell/antibody epitopes ([143] and contributing references) and protease recognition sites ([144] and contributing references).

Modifying peptide libraries before selection

Non-natural amino acid, D-amino acid and more recently structural constraint [145], cyclization [146], and truncation [142] strategies have also been applied to biologically-derived peptide libraries before selection. The advantages of this lie in selection of peptides with already enhanced therapeutic properties, and improved tolerance to future modifications of this nature. In addition, non-natural amino acid substitution and at least partial non-peptidic backbone replacement can be achieved in a non-genetic peptide library selection context.

Expanding the genetic code

Non-natural and D-amino acids have been incorporated into peptides by a number of techniques in the last 25 years [72,147]. Firstly, tRNA can be acylated with non-natural amino acids. This can be achieved by chemical synthesis, aminoacyl tRNA-transferases with mutated substrate recognition or flexizymes, ribozyme aminoacyl tRNA-transferases [148]. Initially, nonsense tRNA molecules were exploited for this purpose;

however, although this can potentially allow incorporation of a wide range of non-natural amino acids at all three stop codons (amber, ochre and opal), incorporation is inefficient at multiple sites and incorporation ratios are not easily controllable due to competition with the natural, preferred tRNA substrates [149]. The problem of competition can be overcome in *in vitro* library systems by using purified recombinant ribosome, release factors and other expression components [150], which then are supplemented with known quantities of mischarged tRNAs and natural nonsense tRNAs are excluded. This technique can additionally be expanded to use sense tRNAs not required in library construction, whose undesirable amino acids are substituted for desired non-natural residues.

Other strategies such as engineering the ribosome to boost recognition of unusually charged tRNA can increase the efficiency of incorporation of non-natural amino acids [151]. Recent realization that different tRNA constant regions have different affinities for the EF-Tu translation elongation factor could be used to ensure that only the highest affinity tRNA constant regions are used for non-natural amino acid incorporation [152]. Alternatively, mutant ribosomes have been selected to incorporate non-natural residues carried by four base pair anticodon tRNA molecules [153].

Flexizyme strategies [148] have been developed in the last 10 years, and make use of ribozymes – 45-nucleotide long RNA molecules – that have evolved (naturally or artificially) to transfer amino acids to tRNA molecules almost irrespectively of the tRNA anticodon. As such flexizymes can be exploited to charge tRNAs with a wide range of exotic amino acids *in vitro*, many more than are possible using natural aminoacyl-tRNA synthetases. With flexizymes, it is possible to incorporate amino acids with non-natural side chains, D-amino acids and N-substituted amino acids. However, recent focus has been on the use of flexizymes to display macrocyclic peptides, desirable based on their cell penetrating properties and high stability. To date, strategies such as thioether, ‘click’ chemistry and lactam linkages have been successfully applied to mRNA/ribosome display libraries using flexizymes [154].

In the future, two fields of development may achieve entirely novel methods by which non-natural amino acid libraries could be generated. The first field relates to that of unnatural nucleic acid (‘Xeno-NA’ or ‘XNA’) base pairs, nucleotide analogs with altered hydrogen-bonding or predominantly hydrophobic pairing interactions, that have the potential to revolutionize genetic engineering [155]. In the future, we could envisage a system that combines currently achievable efficient transcription of XNA [156], with an engineered ribosome or similar molecular machine that could link

these XNA codons to tRNA molecules with unnatural anticodon bases [157] and charged with non-natural amino acids, to allow the novel incorporation of non-natural amino acids into peptides in *in vitro* display systems. Further, the recently reported successful replication of a bacterium harboring genomic XNA [158] suggests that an *in vivo* display system analogous to PCA could also be built using XNAs.

The second field features investigation of non-ribosomal peptide synthetase (NRPS) complexes [159,160] common to bacterial systems that naturally synthesize non-proteinogenic peptides. These complexes similarly could be exploited in the future for non-natural library expression. The advantage of this would be that the diverse range of non-proteinogenic amino acids and their even more diverse range of chemical properties – from non-peptidic backbones (such as polyketide) to non-natural amino acid side chains (such as vinyl-arginine and methyltyrosine) [160], and geometrical properties such as macrocyclization – could be harnessed to confer benefits such as protease resistance, novel spatial geometries and novel binding interactions. However, there are several hurdles to overcome before this option approaches feasibility, the main being that, unlike for ribosomal synthesis where mRNA provides the genetic encoding for amino acid incorporation, NRPS peptide synthesis does not make use of an equivalent genetic material to encode synthesis. Instead, synthesis is directed by NRPS enzyme substrate specificities [161,162], and thus substrate abundances. Currently, basic research has identified numerous NRPSs and their substrate specificities, and altered these to explore even wider application of these enzymes to *in vitro* production of novel peptides [161]. However, a means of linking synthesis to nucleic acid encoding (or an alternate means with which ribosomal and NRPS amino acid incorporation can be combined) will require much further consideration.

For the non-genetic selection of peptide libraries, short regions of chemically synthesized non-natural amino acids can be combined with α -amino acid libraries by chemical conjugation [163]. In this approach, a peptide library contains a protease recognition site as a non-randomized feature, cleavage at which leaves a suitably reactive residue side chain (i.e. cysteine) which can then form a covalent link to an oligomer of non-natural amino acids generated by SPPS. Similarly, the aforementioned techniques of native chemical ligation and ‘click chemistry’ can be used for the purpose of introducing synthesized non-natural peptides into peptide libraries, as could direct insertion of peptide fragments into library peptides in aziridine amide-catalyzed reactions [134]. Future application of these chemistries to peptide libraries for

genotype–phenotype selection would further facilitate selection of pre-modified peptides.

Structural constraint, cyclization & non-peptidic backbones

For structural constraint of *in vitro* displayed libraries, peptides containing cysteine residues can be circularized by alkylating agents with a preference for thiols, for example, bromomethyl benzene derivatives have been used to circularize peptides displayed on phage capsids [145]. The rapid, specific and amenable chemistry of aryl halides ensures quantitative cyclization of large peptide libraries in aqueous solvents and at near-neutral pH.

Macrocyclization of cysteine-containing peptide libraries can be achieved *in vivo* by library selection systems such as SICLOPPS (split-intein circular ligation of peptides and proteins) [146], whereby a cellular ‘internal protein’ (‘intein’) is split, library peptides are inserted between the two intein halves spaced by sulfur-containing linkers and then refolding of the intein initiates an autocatalytic N-to-S acyl shift that results in release of the intact intein and circularization of the library peptide by its linkers.

Finally, chemical conjugation [163] has the potential to combine library peptides with small molecules featuring non-peptidic backbones for non-genetic peptide selection. Alternatively, peptide libraries chemically synthesized on beads (‘OBOC’ or ‘one bead, one compound’ strategies) can be partially transformed to generate polyamine, polyurea, polythiourea or cyclic small molecule-peptide hybrid or peptidomimetic

libraries by reduction, oxidation, alkylation, acylation and peralkylation [164].

Unlike these strategies, HBS and conjugation have only generally been applied to peptides post-selection. For conjugation, this is generally desirable, as carrier molecules could unnecessarily complicate selection and do not significantly alter binding of library selected peptides. However in the future, further development of strategies may allow HBS and further non-natural peptide modifications to be applied to library peptides.

Conclusion & future perspective

In recent years, there has been an explosion in research centered around peptide-based molecules for PPI interference, as the difficulty of targeting PPIs with small molecules has become apparent. A huge variety of peptide modifications are now available, and can improve virtually any aspect of a peptide’s interaction behavior and pharmacokinetics, such that peptide-like specificity can be combined with small molecule-like stability. However, a somewhat trial-and-error approach may be necessary to probe which modifications are most suitable/advantageous. Recent advances allow limited application of these modification strategies to peptide libraries before selection. In the future, new ways to incorporate these peptide modifications into library display systems will accelerate this process, providing high-throughput peptide modification. In particular, further development of expanded codon sets including entirely non-natural ones for non-natural amino acid incorporation into peptide library display could be expected. Sim-

Executive summary

Peptides as therapeutic protein–protein interaction inhibitors

- In recent years, peptides and particularly peptide-derived agents have begun to show considerable promise as molecules with superior selectivity in molecular recognition over traditional small molecules, making excellent protein–protein interaction (PPI) interfering agents.

Peptide libraries as a source of PPI inhibitors

- The advent and development of a broad range of peptide library selection systems has made identifying high-affinity peptide ligands relatively easy, rapid and informative. The strengths of these strategies lie in coverage of vast chemical functionality and spatial orientation, and panning for binding sequences under highly manipulable selection conditions.

Peptide development for therapeutic administration

- Peptides display a previously unappreciated amenability to modification to improve therapeutic potential that matches that of chemical modification of small molecules. A host of backbone and side chain modification strategies can improve binding affinities, bioavailabilities, stabilities and overall efficacies way beyond the capabilities of natural peptides.

Modifying peptide libraries before selection

- More recently, extensive adaptation and development of display strategies enables peptide isolation and modification to be combined to further expedite the process of identifying peptide-derived agents with real potential for therapeutic administration.

Future perspective

- In the future, further developments in display systems, peptide modification and integration of both these techniques should enable even greater harnessing of the potential of peptide-based agents as PPI modulators for therapeutic benefit.

ilarly, exploitation of non-ribosomal peptide synthetases for non-natural peptide display may be achieved. Two remaining challenges to peptide-based therapies outside the fields discussed in this review are the identification and validation of disease cell-specific antigens for cell-specific targeting [165], and the related issue of delivery of therapeutic agents to specific cells. Recent [166] and continuing future developments in these complementary fields will undoubtedly promote major advancements of peptide-based 'drugs' to the clinic [7].

References

Papers of special note have been highlighted as:

• of interest; •• of considerable interest.

- McDowall MD, Scott MS, Barton GJ. Human protein–protein interaction prediction. *BMC Bioinformatics* 11(Suppl. 10), P1 (2010).
- Zhang QC, Petrey D, Deng L *et al.* Structure-based prediction of protein–protein interactions on a genome-wide scale. *Nature* 490(7421), 556–560 (2012).
- Guo WX, Wisniewski JA, Ji HT. Hot spot-based design of small-molecule inhibitors for protein–protein interactions. *Bioorg. Med. Chem. Lett.* 24(11), 2546–2554 (2014).
- Jochim AL, Arora PS. Assessment of helical interfaces in protein–protein interactions. *Mol. Biosyst.* 5(9), 924–926 (2009).
- Jochim AL, Arora PS. Systematic analysis of helical protein interfaces reveals targets for synthetic inhibitors. *ACS Chem. Biol.* 5(10), 919–923 (2010).
- Diao L, Meibohm B. Pharmacokinetics and pharmacokinetic-pharmacodynamic correlations of therapeutic peptides. *Clin. Pharmacokinet.* 52(10), 855–868 (2013).
- Kaspar AA, Reichert JM. Future directions for peptide therapeutics development. *Drug Discov. Today* 18(17–18), 807–817 (2013).
- Odegrip R, Coomber D, Eldridge B *et al.* CIS display: *in vitro* selection of peptides from libraries of protein-DNA complexes. *Proc. Natl Acad. Sci. USA* 101(9), 2806–2810 (2004).
- Griffiths AD, Williams SC, Hartley O *et al.* Isolation of high-affinity human-antibodies directly from large synthetic repertoires. *EMBO J.* 13(14), 3245–3260 (1994).
- Fields S, Song OK. A novel genetic system to detect protein protein interactions. *Nature* 340(6230), 245–246 (1989).
- Ponsel D, Neugebauer J, Laderzki-Baehs K, Tissot K. High affinity, developability and functional size: The holy grail of combinatorial antibody library generation. *Molecules* 16(5), 3675–3700 (2011).
- Mahon CM, Lambert MA, Glanville J *et al.* Comprehensive interrogation of a minimalist synthetic CDR-H3 library and its ability to generate antibodies with therapeutic potential. *J. Mol. Biol.* 425(10), 1712–1730 (2013).
- Crick FH. Codon–anticodon pairing: The wobble hypothesis. *J. Mol. Biol.* 19(2), 548–555 (1966).
- Kammann M, Laufs J, Schell J, Gronenborn B. Rapid insertional mutagenesis of DNA by polymerase chain-reaction (PCR). *Nucleic Acids Res.* 17(13), 5404–5404 (1989).
- Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain-reaction. *Gene* 77(1), 51–59 (1989).
- Virnekas B, Ge LM, Pluckthun A, Schneider KC, Wellnhofer G, Moroney SE. Trinucleotide phosphoramidites - ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis. *Nucleic Acids Res.* 22(25), 5600–5607 (1994).
- Ashraf M, Frigotto L, Smith ME *et al.* ProxiMAX randomization: A new technology for non-degenerate saturation mutagenesis of contiguous codons. *Biochem. Soc. Trans.* 41, 1189–1194 (2013).
- This paper describes a powerful ligation-amplification-digestion technique for construction of randomized libraries without redundancies or bias in encoded amino acids, thus maximizing peptide variant coverage by library display systems to maximize the probability of isolating high-affinity peptides.
- Tang L, Gao H, Zhu X, Wang X, Zhou M, Jiang R. Construction of “small-intelligent” focused mutagenesis libraries using well-designed combinatorial degenerate primers. *Biotechniques* 52(3), 149–158 (2012).
- Janczyk M, Appel B, Springstube D, Fritz HJ, Muller S. A new and convenient approach for the preparation of beta-cyanoethyl protected trinucleotide phosphoramidites. *Org. Biomol. Chem.* 10(8), 1510–1513 (2012).
- Arunachalam TS, Wichert C, Appel B, Muller S. Mixed oligonucleotides for random mutagenesis: Best way of making them. *Org. Biomol. Chem.* 10(24), 4641–4650 (2012).
- Van den Brulle J, Fischer M, Langmann T *et al.* A novel solid phase technology for high-throughput gene synthesis. *Biotechniques* 45(3), 340–343 (2008).
- Hida K, Hanes J, Ostermeier M. Directed evolution for drug and nucleic acid delivery. *Adv. Drug Del. Rev.* 59(15), 1562–1578 (2007).
- Whitney M, Crisp JL, Olson ES *et al.* Parallel *in vivo* and *in vitro* selection using phage display identifies protease-dependent tumor-targeting peptides. *J. Biol. Chem.* 285(29), 22532–22541 (2010).
- Sergeeva A, Kolonin MG, Molldrem JJ, Pasqualini R, Arap W. Display technologies: Application for the discovery of drug and gene delivery agents. *Adv. Drug Del. Rev.* 58(15), 1622–1654 (2006).

Financial & competing interest disclosure

Chris Ullman is CSO at Isogenica Ltd, who provide peptide isolation against therapeutic targets by the 'CIS display' selection system. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

- 25 Ullman CG, Frigotto L, Cooley RN. *In vitro* methods for peptide display and their applications. *Brief. Funct. Genom.* 10(3), 125–134 (2011).
- 26 Wada A. Development of next-generation peptide binders using *in vitro* display technologies and their potential applications. *Front. Immunol.* 4(224), 1–6 (2013).
- 27 Shimizu Y, Kanamori T, Ueda T. Protein synthesis by pure translation systems. *Methods* 36(3), 299–304 (2005).
- 28 Doi N, Takashima H, Wada A, Oishi Y, Nagano T, Yanagawa H. Photocleavable linkage between genotype and phenotype for rapid and efficient recovery of nucleic acids encoding affinity-selected proteins. *J. Biotechnol.* 131(3), 231–239 (2007).
- 29 Mattheakis LC, Bhatt RR, Dower WJ. An *in vitro* polysome display system for identifying ligands from very large peptide libraries. *Proc. Natl Acad. Sci. USA* 91(19), 9022–9026 (1994).
- 30 Masai H, Kaziro Y, Arai KI. Definition of OriR, the minimum DNA segment essential for initiation of R1 plasmid replication *in vitro*. *Proc. Natl Acad. Sci. USA* 80(22), 6814–6818 (1983).
- 31 Eldridge B, Cooley RN, Odegrip R, McGregor DP, FitzGerald KJ, Ullman CG. An *in vitro* selection strategy for conferring protease resistance to ligand binding peptides. *Protein Eng. Des. Sel.* 22(11), 691–698 (2009).
- 32 Reiersen H, Lobersli I, Loset GA *et al.* Covalent antibody display - an *in vitro* antibody-DNA library selection system. *Nucleic Acids Res.* 33(1), e10 (2005).
- 33 Bertschinger J, Neri D. Covalent DNA display as a novel tool for directed evolution of proteins *in vitro*. *Protein Eng. Des. Sel.* 17(9), 699–707 (2004).
- 34 Roberts RW, Szostak JW. RNA-peptide fusions for the *in vitro* selection of peptides and proteins. *Proc. Natl Acad. Sci. USA* 94(23), 12297–12302 (1997).
- 35 Tawfik DS, Griffiths AD. Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.* 16(7), 652–656 (1998).
- 36 Tawfik DS, Griffiths AD. Directed evolution of an extremely fast phosphotriesterase by *in vitro* compartmentalization. *EMBO J.* 22(1), 24–35 (2003).
- 37 Smith GP. Filamentous fusion phage – novel expression vectors that display cloned antigens on the virion surface. *Science* 228(4705), 1315–1317 (1985).
- 38 Doerner A, Rhiel L, Zielonka S, Kolmar H. Therapeutic antibody engineering by high efficiency cell screening. *FEBS Lett.* 588(2), 278–287 (2014).
- 39 Makela AR, Oker-Blom C. The baculovirus display technology – an evolving instrument for molecular screening and drug delivery. *Comb. Chem. High Throughput Screen.* 11(2), 86–98 (2008).
- 40 Pelletier JN, Arndt KM, Pluckthun A, Michnick SW. An *in vivo* library-versus-library selection of optimized protein–protein interactions. *Nat. Biotechnol.* 17(7), 683–690 (1999).
- 41 Hu JC, O'Shea EK, Kim PS, Sauer RT. Sequence requirements for coiled-coils – analysis with lambda-repressor-GCN4 leucine zipper fusions. *Science* 250(4986), 1400–1403 (1990).
- 42 Pelletier JN, Campbell-Valois FX, Michnick SW. Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proc. Natl Acad. Sci. USA* 95(21), 12141–12146 (1998).
- 43 Kerppola TK. Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. *Nat. Protoc.* 1(3), 1278–1286 (2006).
- 44 Smith GP, Petrenko VA. Phage display. *Chem. Rev.* 97(2), 391–410 (1997).
- 45 Arap MA. Phage display technology – applications and innovations. *Genet. Mol. Biol.* 28(1), 1–9 (2005).
- 46 Beghetto E, Gargano N. Lambda-display: a powerful tool for antigen discovery. *Molecules* 16(4), 3089–3105 (2011).
- 47 Pasqualini R, Ruoslahti E. Organ targeting *in vivo* using phage display peptide libraries. *Nature* 380(6572), 364–366 (1996).
- Represents one of the first studies in which phage display was shown to be capable of enrichment of peptides specific to extracellular targets in an *in vivo* setting, thus representing a superior embodiment of *in vitro* phage display for concomitant selection of peptide systemic stability.
- 48 Duerr DM, White SJ, Schluesener HJ. Identification of peptide sequences that induce the transport of phage across the gastrointestinal mucosal barrier. *J. Virol. Methods* 116(2), 177–180 (2004).
- 49 Baek DS, Kim YS. Construction of a large synthetic human Fab antibody library on yeast cell surface by optimized yeast mating. *J. Microbiol. Biotechnol.* 24(3), 408–420 (2014).
- 50 Mason JM, Schmitz MA, Mueller K, Arndt KM. Semirational design of Jun-Fos coiled coils with increased affinity: universal implications for leucine zipper prediction and design. *Proc. Natl Acad. Sci. USA* 103(24), 8989–8994 (2006).
- 51 Mason JM, Muller KM, Arndt KM. Positive aspects of negative design: simultaneous selection of specificity and interaction stability. *Biochemistry* 46(16), 4804–4814 (2007).
- 52 Mason JM. Design and development of peptides and peptide mimetics as antagonists for therapeutic intervention. *Future Med. Chem.* 2(12), 1813–1822 (2010).
- 53 Nieuwlandt D. *In vitro* selection of functional nucleic acid sequences. *Curr. Issues Mol. Biol.* 2(1), 9–16 (2000).
- 54 Remy I, Michnick SW. Clonal selection and *in vivo* quantitation of protein interactions with protein-fragment complementation assays. *Proc. Natl Acad. Sci. USA* 96(10), 5394–5399 (1999).
- 55 Bentley DR, Balasubramanian S, Swerdlow HP *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456(7218), 53–59 (2008).
- 56 Rothberg JM, Hinz W, Rearick TM *et al.* An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475(7356), 348–352 (2011).
- 57 Eid J, Fehr A, Gray J *et al.* Real-time DNA sequencing from single polymerase molecules. *Science* 323(5910), 133–138 (2009).
- 58 Howorka S, Cheley S, Bayley H. Sequence-specific detection of individual DNA strands using engineered nanopores. *Nat. Biotechnol.* 19(7), 636–639 (2001).

- 59 Ravn U, Gueneau F, Baerlocher L *et al.* By-passing *in vitro* screening-next generation sequencing technologies applied to antibody display and in silico candidate selection. *Nucleic Acids Res.* 38(21), e193 (2010).
- 60 Ruff AJ, Dennig A, Schwaneberg U. To get what we aim for: progress in diversity generation methods. *FEBS J.* 280(13), 2961–2978 (2013).
- 61 Esvelt KM, Carlson JC, Liu DR. A system for the continuous directed evolution of biomolecules. *Nature* 472(7344), 499–503 (2011).
- 62 Carlson JC, Badran AH, Guggiana-Nilo DA, Liu DR. Negative selection and stringency modulation in phage-assisted continuous evolution. *Nat. Chem. Biol.* 10(3), 216–222 (2014).
- 63 Beyer M, Nesterov A, Block I *et al.* Combinatorial synthesis of peptide arrays onto a microchip. *Science* 318(5858), 1888–1888 (2007).
- 64 Stadler V, Felgenhauer T, Beyer M *et al.* Combinatorial synthesis of peptide arrays with a laser printer. *Angew. Chem. Int. Ed.* 47(37), 7132–7135 (2008).
- 65 Buus S, Rockberg J, Forsstrom B, Nilsson P, Uhlen M, Schafer-Nielsen C. High-resolution mapping of linear antibody epitopes using ultrahigh-density peptide microarrays. *Mol. Cell. Proteomics* 11(12), 1790–1800 (2012).
- 66 Craik DJ, Fairlie DP, Liras S, Price D. The future of peptide-based drugs. *Chem. Biol. Drug Des.* 81(1), 136–147 (2013).
- 67 Bryson CJ, Jones TD, Baker MP. Prediction of immunogenicity of therapeutic proteins: validity of computational tools. *Biodrugs* 24(1), 1–8 (2010).
- 68 Kilby JM, Hopkins S, Venetta TM *et al.* Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. *Nat. Med.* 4(11), 1302–1307 (1998).
- 69 Reichert J. Development trends for peptide therapeutics. www.peptidetherapeutics.org/PTF_report_summary_2010.pdf
- 70 Liskamp RMJ, Rijkers DTS, Kruijtz JAW, Kemmink J. Peptides and proteins as a continuing exciting source of inspiration for peptidomimetics. *ChemBioChem* 12(11), 1626–1653 (2011).
- **Highlights the main attractions of peptides for development into therapeutic agents, as a combination of their broad chemical functionality, range of secondary structures and amenability to modification, with a focus on the transformations available for conversion of unmodified peptides into stable and cell penetrating molecules while retaining their high target specificity.**
- 71 Adessi C, Soto C. Converting a peptide into a drug: strategies to improve stability and bioavailability. *Curr. Med. Chem.* 9(9), 963–978 (2002).
- 72 Young TS, Schultz PG. Beyond the canonical 20 amino acids: expanding the genetic lexicon. *J. Biol. Chem.* 285(15), 11039–11044 (2010).
- 73 Cheng RP, Gellman SH, DeGrado WF. Beta-peptides: from structure to function. *Chem. Rev.* 101(10), 3219–3232 (2001).
- 74 Simon RJ, Kania RS, Zuckermann RN *et al.* Peptoids – a modular approach to drug discovery. *Proc. Natl Acad. Sci. USA* 89(20), 9367–9371 (1992).
- 75 Tang Y, Tirrell DA. Biosynthesis of a highly stable coiled-coil protein containing hexafluoroleucine in an engineered bacterial host. *J. Am. Chem. Soc.* 123(44), 11089–11090 (2001).
- 76 Seebach D, Overhand M, Kuhnle FNM *et al.* Beta-peptides: synthesis by Arndt-Eistert homologation with concomitant peptide coupling. Structure determination by NMR and CD spectroscopy and by X-ray crystallography. Helical secondary structure of a beta-hexapeptide in solution and its stability towards pepsin. *Helv. Chim. Acta* 79(4), 913–941 (1996).
- 77 Maurer PH. Antigenicity of polypeptides (poly alpha amino acids). 13. Immunological studies with synthetic polymers containing only D- or D- and L-alpha-amino acids. *J. Exp. Med.* 121, (1965).
- 78 Moree WJ, Vandermael GA, Liskamp RMJ. Synthesis of peptidosulfonamides and peptidosulfonamides – peptidomimetics containing the sulfonamide or sulfonamide transition-state isostere. *J. Org. Chem.* 60(16), 5157–5169 (1995).
- 79 Garner J, Harding MM. Design and synthesis of alpha-helical peptides and mimetics. *Org. Biomol. Chem.* 5(22), 3577–3585 (2007).
- 80 Sutherell CL, Thompson S, Scott RTW, Hamilton AD. Aryl-linked imidazolidin-2-ones as non-peptidic beta-strand mimetics. *Chem. Commun.* 48(79), 9834–9836 (2012).
- 81 Johnson LM, Gellman SH. Alpha-helix mimicry with alpha/beta-peptides. *Methods Enzymol.* 523, 407–429 (2013).
- 82 Dewal MB, Firestone SM. Non-peptidic alpha-helical mimetics as protein–protein interaction inhibitors. *Curr. Med. Chem.* 18(16), 2420–2428 (2011).
- 83 Adler MJ, Hamilton AD. Oligophenylaminones as scaffolds for alpha-helix mimicry. *J. Org. Chem.* 76(17), 7040–7047 (2011).
- 84 Prabhakaran P, Barnard A, Murphy NS, Kilner CA, Edwards TA, Wilson AJ. Aromatic oligoamide foldamers with a “wet edge” as inhibitors of the alpha-helix-mediated p53-Hdm2 protein–protein interaction. *Eur. J. Org. Chem.* (17), 3504–3512 (2013).
- 85 Ross NT, Katt WP, Hamilton AD. Synthetic mimetics of protein secondary structure domains. *Philos. Trans. A Math. Phys. Eng. Sci.* 368(1914), 989–1008 (2010).
- 86 Goodman M, Chorev M. Concept of linear modified retro-peptide structures. *Acc. Chem. Res.* 12(1), 1–7 (1979).
- 87 Acerra N, Kad NM, Griffith DA, Ott S, Crowther DC, Mason JM. Retro-inverso of intracellular selected beta-amyloid-interacting peptides: implications for a novel Alzheimer’s disease treatment. *Biochemistry* 53(13), 2101–2111 (2014).
- 88 Yamada R, Kera Y. D-amino acid hydrolysing enzymes. *EXS* 85, 145–155 (1998).
- 89 Akita H, Suzuki H, Doi K, Ohshima T. Efficient synthesis of D-branched-chain amino acids and their labeled compounds with stable isotopes using D-amino acid dehydrogenase. *Appl. Microbiol. Biotechnol.* 98(3), 1135–1143 (2014).

- 90 Schumacher TNM, Mayr LM, Minor DL, Milhollen MA, Burgess MW, Kim PS. Identification of D-peptide ligands through mirror-image phage display. *Science* 271(5257), 1854–1857 (1996).
- 91 Lifson S, Roig A. On the theory of helix-coil transitions in polypeptides. *J. Chem. Phys.* 34, 1963–1974 (1961).
- 92 Rao T, Ruiz-Gómez G, Hill TA, Hoang HN, Fairlie DP, Mason JM. Truncated and helix-constrained peptides with high affinity and specificity for the cFos coiled-coil of AP-1. *PLoS ONE* 8(3), e59415–e59426 (2013).
- 93 Yin H. Constrained peptides as miniature protein structures. *ISRN Biochem.* 2012, 1–15 (2012).
- 94 Hill TA, Shepherd NE, Diness F, Fairlie DP. Constraining cyclic peptides to mimic protein structure motifs. *Angew. Chem. Int. Ed.* 53(48), 13020–13041 (2014).
- 95 de Araujo AD, Hoang HN, Kok WM *et al.* Comparative alpha-helicity of cyclic pentapeptides in water. *Angew. Chem. Int. Ed.* 53(27), 6965–6969 (2014).
- 96 Harrison RS, Shepherd NE, Hoang HN *et al.* Downsizing human, bacterial, and viral proteins to short water-stable alpha helices that maintain biological potency. *Proc. Natl Acad. Sci. USA* 107(26), 11686–11691 (2010).
- 97 Verdine GL, Hilinski GJ. Stapled peptides for intracellular drug targets. In: *Methods in Enzymology: Protein Engineering for Therapeutics, Volume 203, Pt B*. Wittrup KD, Verdine GL (Eds). Elsevier Academic Press Inc, San Diego, CA, USA, 3–33 (2012).
- 98 Shepherd NE, Hoang HN, Abbenante G, Fairlie DP. Single turn peptide alpha helices with exceptional stability in water. *J. Am. Chem. Soc.* 127(9), 2974–2983 (2005).
- 99 Eferl R, Wagner EF. AP-1: a double-edged sword in tumorigenesis. *Nat. Rev. Cancer* 3(11), 859–868 (2003).
- 100 Bernal F, Wade M, Godes M *et al.* A stapled p53 helix overcomes HdmX-mediated suppression of p53. *Cancer Cell* 18(5), 411–422 (2010).
- 101 Baek S, Kutchukian PS, Verdine GL *et al.* Structure of the stapled p53 peptide bound to Mdm2. *J. Am. Chem. Soc.* 134(1), 103–106 (2012).
- 102 Walensky LD, Kung AL, Escher I *et al.* Activation of apoptosis *in vivo* by a hydrocarbon-stapled BH3 helix. *Science* 305(5689), 1466–1470 (2004).
- 103 Sun T-L, Sun Y, Lee C-C, Huang HW. Membrane permeability of hydrocarbon-cross-linked peptides. *Biophys. J.* 104(9), 1923–1932 (2013).
- 104 Bird GH, Madani N, Perry AF *et al.* Hydrocarbon double-stapling remedies the proteolytic instability of a lengthy peptide therapeutic. *Proc. Natl Acad. Sci. USA* 107(32), 14093–14098 (2010).
- 105 Fairlie DP, Shepherd NE: WO2005090388A1(2005).
- 106 Timmerman P, Beld J, Puijk WC, Meloen RH. Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces. *ChemBioChem* 6(5), 821–824 (2005).
- 107 Jo H, Meinhardt N, Wu Y *et al.* Development of alpha-helical calpain probes by mimicking a natural protein–protein interaction. *J. Am. Chem. Soc.* 134(42), 17704–17713 (2012).
- 108 Schafmeister CE, Po J, Verdine GL. An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. *J. Am. Chem. Soc.* 122(24), 5891–5892 (2000).
- 109 Kelso MJ, Beyer RL, Hoang HN *et al.* Alpha-turn mimetics: short peptide alpha-helices composed of cyclic metallopentapeptide modules. *J. Am. Chem. Soc.* 126(15), 4828–4842 (2004).
- 110 Grigoryev Y. Stapled peptide to enter human testing, but affinity questions remain. *Nat. Med.* 19(2), 120–120 (2013).
- 111 Chapman RN, Dimartino G, Arora PS. A highly stable short alpha-helix constrained by a main-chain hydrogen-bond surrogate. *J. Am. Chem. Soc.* 126(39), 12252–12253 (2004).
- 112 Patgiri A, Jochim AL, Arora PS. A hydrogen bond surrogate approach for stabilization of short peptide sequences in alpha-helical conformation. *Acc. Chem. Res.* 41(10), 1289–1300 (2008).
- 113 Kemp DS, Curran TP, Davis WM, Boyd JG, Muendel C. Studies of N-terminal templates for alpha-helix formation – synthesis and conformational analysis of (2S, 5S, 8S, 11S)-1-acteyl-1,4-diaza-3-keto-5-carboxy-10-thiatricyclo[2.8.1.0^{4,8}] tridecane (Ac-Hel1-OH). *J. Org. Chem.* 56(23), 6672–6682 (1991).
- 114 Schneider JP, Kelly JW. Templates that induce alpha-helical, beta-sheet, and loop conformations. *Chem. Rev.* 95(6), 2169–2187 (1995).
- 115 Loughlin WA, Tyndall JDA, Glenn MP, Hill TA, Fairlie DP. Update 1 of: beta-strand mimetics. *Chem. Rev.* 110(6), (2010).
- 116 Weidle UH, Auer J, Brinkmann U, Georges G, Tiefenthaler G. The emerging role of new protein scaffold-based agents for treatment of cancer. *Cancer Genom. Proteom.* 10(4), (2013).
- 117 de Vega MJP, Garcia-Aranda MI, Gonzalez-Muniz R. A role for ring-closing metathesis in medicinal chemistry: mimicking secondary architectures in bioactive peptides. *Med. Res. Rev.* 31(5), 677–715 (2011).
- 118 Kemp DS, Bowen BR. Synthesis of peptide-functionalized diacylaminoepindolidiones as templates for beta-sheet formation. *Tetrahedron Lett.* 29(40), 5077–5080 (1988).
- 119 Fuller AA, Du D, Liu F *et al.* Evaluating beta-turn mimics as beta-sheet folding nucleators. *Proc. Natl Acad. Sci. USA* 106(27), 11067–11072 (2009).
- 120 Fetrow JS. Protein motifs.6. Omega-loops – nonregular secondary structures significant in protein function and stability. *FASEB J.* 9(9), 708–717 (1995).
- 121 Sarabu R, Lovey K, Madison VS *et al.* Design, synthesis, and 3-dimensional structural characterization of a constrained omega-loop excised from interleukin-1-alpha. *Tetrahedron* 49(17), 3629–3640 (1993).
- 122 Madala PK, Tyndall JDA, Nall T, Fairlie DP. Update 1 of: Proteases universally recognize beta strands in their active sites. *Chem. Rev.* 110(6), PR1–PR31 (2010).
- 123 Bockus AT, McEwen CM, Lokey RS. Form and function in cyclic peptide natural products: a pharmacokinetic perspective. *Curr. Top. Med. Chem.* 13(7), 821–836 (2013).
- 124 White CJ, Yudin AK. Contemporary strategies for peptide macrocyclization. *Nat. Chem.* 3(7), 509–524 (2011).

- Recent in-depth review of metal and sulfur-catalyzed cyclizations, ring-closing and ring-contraction strategies, azide-alkyne additions (“click” chemistry), electrostatically controlled cyclizations (exploiting peptide conformational preference for cyclization-prone conformers in organic polar solvents) and multicomponent reaction cyclizations (featuring modification of multiple moieties simultaneously to expedite formation of a cyclization-ready conformer), as well as structural and entropic considerations, for efficient macrocyclization of peptides during/following synthesis or in a library display context.
- 125 Rezai T, Bock JE, Zhou MV, Kalyanaraman C, Lokey RS, Jacobson MP. Conformational flexibility, internal hydrogen bonding, and passive membrane permeability: successful in silico prediction of the relative permeabilities of cyclic peptides. *J. Am. Chem. Soc.* 128(43), 14073–14080 (2006).
- 126 White TR, Renzelman CM, Rand AC *et al.* On-resin N-methylation of cyclic peptides for discovery of orally bioavailable scaffolds. *Nat. Chem. Biol.* 7(11), 810–817 (2011).
- 127 Bruns C, Lewis I, Briner U, Meno-Tetang G, Weckbecker G. Som230: A novel somatostatin peptidomimetic with broad somatotropin release inhibiting factor (SRIF) receptor binding and a unique antiseecretory profile. *Eur. J. Endocrinol.* 146(5), 707–716 (2002).
- 128 Colao A, Petersenn S, Newell-Price J *et al.* A 12-month phase 3 study of pasireotide in Cushing’s disease. *N. Engl. J. Med.* 366(10), 914–924 (2012).
- 129 Corey GR, Kabler H, Mehra P *et al.* Single-dose oritavancin in the treatment of acute bacterial skin infections. *N. Engl. J. Med.* 370(23), 2180–2190 (2014).
- 130 Turner RA, Oliver AG, Lokey RS. Click chemistry as a macrocyclization tool in the solid-phase synthesis of small cyclic peptides. *Org. Lett.* 9(24), 5011–5014 (2007).
- 131 Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. Synthesis of proteins by native chemical ligation. *Science* 266(5186), 776–779 (1994).
- 132 Kimura R, Camarero JA. Expressed protein ligation: a new tool for the biosynthesis of cyclic polypeptides. *Protein Pept. Lett.* 12(8), 789–794 (2005).
- 133 van de Langemheen H, van Ufford HCQ, Kruijtz JAW, Liskamp RMJ. Efficient synthesis of protein mimics by sequential native chemical ligation. *Org. Lett.* 16(8), 2138–2141 (2014).
- 134 White CJ, Hickey JL, Scully CCG, Yudin AK. Site-specific integration of amino acid fragments into cyclic peptides. *J. Am. Chem. Soc.* 136(10), 3728–3731 (2014).
- 135 Choi YS, Lee JY, Suh JS *et al.* Cell penetrating peptides for tumor targeting. *Curr. Pharm. Biotechnol.* 12(8), 1166–1182 (2011).
- 136 Fawell S, Seery J, Daikh Y *et al.* Tat-mediated delivery of heterologous proteins into cells. *Proc. Natl Acad. Sci. USA* 91(2), 664–668 (1994).
- 137 Madani F, Lindberg S, Langel U, Futaki S, Graslund A. Mechanisms of cellular uptake of cell-penetrating peptides. *J. Biophys.* 2011, 414729–414729 (2011).
- 138 Harris JM, Chess RB. Effect of pegylation on pharmaceuticals. *Nat. Rev. Drug Discov.* 2(3), 214–221 (2003).
- 139 Park M, Jardetzky TS, Barron AE. NMEGylation: a novel modification to enhance the bioavailability of therapeutic peptides. *Biopolymers* 96(5), 688–693 (2011).
- 140 Roberts MJ, Bentley MD, Harris JM. Chemistry for peptide and protein PEGylation. *Adv. Drug Del. Rev.* 54(4), 459–476 (2002).
- 141 Lu YA, Felix AM. Pegylated peptides.2. Solid-phase synthesis of amino-chain, carboxy-chain and side-chain pegylated peptides. *Int. J. Pept. Protein Res.* 43(2), 127–138 (1994).
- 142 Crooks RO, Rao T, Mason JM. Truncation, randomization, and selection: generation of a reduced length c-Jun antagonist that retains high interaction stability. *J. Biol. Chem.* 286(34), 29470–29479 (2011).
- 143 Vita R, Zarebski L, Greenbaum JA *et al.* The Immune Epitope Database 2.0. *Nucleic Acids Res.* 38 D854–D862 (2010).
- 144 Gasteiger E, Hoogland C, Gattiker A *et al.* Protein identification and analysis tools on the ExPASy server. In: *The Proteomics Protocols Handbook*, Walker JM (Ed.). Humana Press, Totowa, NJ, USA (2005).
- 145 Heinis C, Rutherford T, Freund S, Winter G. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat. Chem. Biol.* 5(7), 502–507 (2009).
- Describes an ingenious approach for display of cyclic peptide libraries that features post-expression cyclization of linear peptide cysteine thiols by bromobenzene derivatives to allow selection of already-modified peptides and thus avoid intolerance issues associated with post-selection cyclization.
- 146 Tavassoli A, Benkovic SJ. Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in E-coli. *Nat. Protoc.* 2(5), 1126–1133 (2007).
- 147 Kawakami T, Murakami H. Genetically encoded libraries of nonstandard peptides. *J. Nucl. Acids* 2012, (2012).
- 148 Goto Y, Katoh T, Suga H. Flexizymes for genetic code reprogramming. *Nat. Protoc.* 6(6), 779–790 (2011).
- Details protocols for the charging of flexizyme ribozymes with non-natural amino acids and their incorporation into non-natural peptides in *in vitro* transcription-translation reactions for use in library display systems.
- 149 Noren CJ, Anthony-Cahill SJ, Griffith MC, Schultz PG. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* 244(4901), 182–188 (1989).
- 150 Shimizu Y, Inoue A, Tomari Y *et al.* Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19(8), 751–755 (2001).
- 151 Wang L, Schultz PG. Expanding the genetic code. *Angew. Chem. Int. Ed.* 44(1), 34–66 (2005).
- 152 Jeong KW, Pavlov MY, Kwiatkowski M, Ehrenberg M, Forster AC. A tRNA body with high affinity for EF-Tu hastens ribosomal incorporation of unnatural amino acids. *RNA* 20(5), 632–643 (2014).
- 153 Neumann H, Wang KH, Davis L, Garcia-Alai M, Chin JW. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464(7287), 441–444 (2010).

- 154 Ito K, Passioura T, Suga H. Technologies for the synthesis of mRNA-encoding libraries and discovery of bioactive natural product-inspired non-traditional macrocyclic peptides. *Molecules* 18(3), 3502–3528 (2013).
- 155 Pinheiro VB, Holliger P. The XNA world: progress towards replication and evolution of synthetic genetic polymers. *Curr. Opin. Chem. Biol.* 16(3–4), 245–252 (2012).
- 156 Seo YJ, Matsuda S, Romesberg FE. Transcription of an expanded genetic alphabet. *J. Am. Chem. Soc.* 131(14), 5046–5047 (2009).
- This paper was one of the first to describe efficient transcription of unnatural nucleic acid bases based on hydrophobic pairing by unmodified RNA polymerase with a fidelity equivalent to that of natural DNA bases, demonstrating the potential of this approach to in the future be used for wider applications.
- 157 Fukunaga R, Harada Y, Hirao I, Yokoyama S. Phosphoserine aminoacylation of tRNA bearing an unnatural base anticodon. *Biochem. Biophys. Res. Commun.* 372(3), 480–485 (2008).
- 158 Malyshev DA, Dhami K, Lavergne T *et al.* A semi-synthetic organism with an expanded genetic alphabet. *Nature* 509(7500), 385–388 (2014).
- 159 Marahiel MA. Working outside the protein-synthesis rules: Insights into non-ribosomal peptide synthesis. *J. Pept. Sci.* 15(12), 799–807 (2009).
- 160 Walsh CT, Brien RVO, Khosla C. Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polyketide scaffolds. *Angew. Chem. Int. Ed.* 52(28), 7098–7124 (2013).
- 161 Villiers B, Hollfelder F. Directed evolution of a gatekeeper domain in nonribosomal peptide synthesis. *Chem. Biol.* 18(10), 1290–1299 (2011).
- Describes the successful evolution of altered substrate specificity in the substrate-accepting adenylation domain of a nonribosomal peptide synthetase as an alternative approach to achieve non-natural peptide synthesis.
- 162 Lautru S, Challis GL. Substrate recognition by nonribosomal peptide synthetase multi-enzymes. *Microbiology* 150, 1629–1636 (2004).
- 163 Lee BC, Zuckermann RN. Protein side-chain translocation mutagenesis via incorporation of peptoid residues. *ACS Chem. Biol.* 6(12), 1367–1374 (2011).
- 164 Kappel JC, Fan YC, Lam KS. Global transformation of OBOC combinatorial peptide libraries into OBOC polyamine and small molecule libraries. *J. Comb. Chem.* 10(2), 333–342 (2008).
- 165 Hinrichs CS, Restifo NP. Reassessing target antigens for adoptive T-cell therapy. *Nat. Biotechnol.* 31(11), 999–1008 (2013).
- 166 Jin SE, Jin HE, Hong SS. Targeted delivery system of nanobiomaterials in anticancer therapy: From cells to clinics. *BioMed Res. Int.* doi:10.1155/2014/814208 (2014) (Epub ahead of print).