Exploiting Overlapping Advantages of in vitro and in cellulo Selection Systems to Isolate a Novel High-affinity cJun Antagonist

Daniel Baxter\textsuperscript{1,2}, Christopher G. Ullman\textsuperscript{2}, Laura Frigotto\textsuperscript{2}, and Jody M. Mason\textsuperscript{1}

\textsuperscript{1}Dept of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY

\textsuperscript{2}Isogenica Ltd, Chesterford Research Park, Little Chesterford, Essex CB10 1XL UK

\textsuperscript{1}To whom correspondence should be addressed: j.mason@bath.ac.uk

Running title: Combining In vitro and In cellulo library screening

*Corresponding author: Dr Jody M. Mason (E: j.mason@bath.ac.uk; T: +441225386867)

Keywords: Coiled coil; Activator Protein-1; transcription factor; Peptide antagonist
ABSTRACT

We have combined two peptide library-screening systems, exploiting the benefits offered by both to select novel antagonistic agents of cJun. CIS display is an in vitro cell-free system that allows very large libraries ($\leq 10^{14}$) to be interrogated. However, affinity-based screening conditions can poorly reflect those relevant to therapeutic application, particularly for difficult intracellular targets, and can lead to false positives. In contrast, an in cellulo screening system such as the Protein-fragment Complementation Assay (PCA) selects peptides with high target affinity while additionally profiling for target-specificity, protease resistance, solubility and lack of toxicity in a more relevant context. A disadvantage is the necessity to transform cells, limiting library sizes that can be screened to $\leq 10^6$. However, by combining both cell-free and cell-based systems, we isolated a peptide (CPW) from a $\sim 10^{10}$ member library, which forms a highly stable interaction with cJun ($T_m$ 63 °C, $K_d$ 750 nM, $\Delta G = -8.2$ kcal/mol) using the oncogenic transcriptional regulator AP-1 as our exemplar target. In contrast, CIS display alone, selected a peptide with low affinity for cJun ($T_m$ 34 °C, $K_d$ 25 μM, $\Delta G = -6.2$ kcal/mol), highlighting the benefit of CIS→PCA. Furthermore, increased library size with CIS→PCA allows the freedom to introduce non-canonical options, such as interfacial aromatics, and solvent exposed options that may allow the molecule to explore alternative structures and interact with greater affinity and efficacy with the target. CIS→PCA therefore offers significant potential as a peptide-library screening platform by synergistically combining the relative attributes of either assay to generate therapeutically interesting compounds that may otherwise not be identified.
INTRODUCTION

Peptide libraries represent a valuable source of potential inhibitors for disease-related protein-protein interactions (PPIs), such as the oncogenic basic-region zipper (bZIP) Activator Protein-1 (AP-1) system\(^1\textsuperscript{-3}\), where smaller more drug-like molecules lack the requisite number of interacting moieties to be effective PPI modulators. Randomised peptide libraries offer huge functional diversity with respect to the range and combination of amino acid side chain physicochemical properties, interactions, and structures that can be achieved. Library screening also has the advantage over rational antagonist design in that affinity for the target is not dependent on comprehensive \textit{a priori} knowledge of the most favourable peptide-target interactions\(^4\).

To antagonise AP-1, a small number of direct inhibitory peptides have been previously derived. Earlier efforts generated peptides and proteins ranging from 46 to approximately 300 residues through native component protein truncation/mutation\(^5\textsuperscript{-10}\). Later, \textit{Protein-fragment Complementation Assay} (PCA)\(^11\) and phage display\(^12\) screening of small libraries (62,000 – 140,000 peptides) generated 37 residue antagonists. More recently, a 42 residue antagonist of the cJun component of AP-1\(^1\) has been designed based on a scoring algorithm derived from peptide interaction microarray data\(^13\). Collectively, these peptides have demonstrated antagonism of AP-1 formation \textit{in vitro}, and in several cases inhibition of oncogenic transformation when transfected and expressed\(^7\textsuperscript{,}10\). However, the higher molecular weight antagonists risk suffering from many of the common shortcomings that prevent therapeutic administration\(^14\textsuperscript{,}15\). Peptides benefit by containing a reduced number of protease recognition sites and immune epitopes. In addition, they are more economical to produce and are easier to modify for cell penetration\(^16\textsuperscript{,}17\). Thus, derivation of short peptidic AP-1 antagonists is an attractive approach towards the development of peptide-based modulatory agents for cancer therapy.

To create antagonists, known peptide binders can be truncated; however, even for parent molecules that have high affinity for the target such as the nanomolar FosW–cJun interaction\(^11\), binding affinity and structural propensity can become rapidly lost. An alternative strategy is to reduce entropic penalty to binding through helix constraint
or iteratively optimise the primary sequence of peptides to compensate for the loss of interaction points with the target. This goal is most easily and effectively achieved using expansive library selection systems.

To derive antagonists of high target affinity and favourable therapeutic-like properties, in cellulo library selection systems have the advantage of screening for target binders in an environment that mimics that in which therapeutic agents are required to function. However, such systems suffer from the requirement for transformation/transfection of cells with library-encoded DNA. This limits the library diversity that can be screened \(^4\). In contrast, cell-free in vitro systems select peptides under less stringent conditions that do not replicate those within the cell, but have a major advantage in their ability to screen much larger peptide libraries to allow much greater exploration of physicochemical, functional and structural peptide space. Another in vitro selection limitation is the possibility that secondary selection pressures will enrich peptides that bind non-specifically to other components in the system, such as magnetic beads, streptavidin, biotin or polypropylene plates\(^20\).

In contrast, in cellulo systems such as bacterial PCA (e.g. based on reactivation of murine DHFR (mDHFR) upon library–target interaction) has the benefit of selecting peptides for therapeutic properties such as target-specificity, aggregation- and protease-resistance, cytosolic stability and non-toxicity \(^21\). In addition, PCA is rapidly and easily performed due to the robustness of bacterial growth. However, the diversities of libraries in PCA-based selection are usually restricted to \(\leq 10^6\) peptides \(^11\). Of the in vitro systems, CIS display \(^22\) combines facile, rapid screening using highly manipulable selection pressures and coverage of similar theoretical library diversities \((\leq 10^{14}\) peptides) \(^23\) to mRNA display \(^24\), ribosome display \(^25\), or phage display alternatives \(^26\). Furthermore, CIS display is based on DNA-encoding of libraries, and so avoids stability issues with mRNA-encoding of libraries in mRNA and ribosome displays.

Previously, cell-based and in vitro methods have been used to screen peptide libraries. However, it was unknown whether such systems could be complementary. The distinct overlap between their advantages and disadvantages led to the hypothesis that successful combination of in vitro and in vivo systems could exploit the benefits of
both to isolate high affinity peptides from hugely diverse libraries with concomitant intracellular refinement of drug-like properties (Figure 1).

A PCA-derived 37 residue cJun antagonist known as FosW \textsuperscript{11} was previously truncated to 32 residues, with PCA further used to re-optimise affinity. The resulting antagonist, 4hFosW, displayed a binding affinity ($\Delta G$) within 11 % of FosW whilst being 13 % shorter (45 nM to 480 nM) \textsuperscript{27}. In contrast, the loss of binding affinity by simple truncation of the FosW sequence by the same amount was measured to be 43 % \textsuperscript{19}. This supports the hypothesis that library selection for more optimal residues to could be successful. Furthermore, by utilising CIS-display to screen a much larger library ($\leq 10^{14}$ peptides) than used to derive 4hFosW (49,152 peptides), it was anticipated that even higher affinities might be possible despite further truncation. Here we describe the novel combination of CIS display and PCA screening systems (“CIS→PCA”) with the aim of creating a more powerful system than either approach alone. In doing so the major aim was to exploit synergy between the two systems, and ultimately generate therapeutically interesting peptides capable of modulating PPIs in therapeutically relevant settings.

RESULTS

Library design

A highly diverse mixed-length peptide library was constructed of various lengths through a modular synthesis of DNA corresponding to individual heptad repeats ($a$-$b$-$c$-$d$-$e$-$f$-$g$). This was expected to represent a coiled coil (CC) library 3.5-heptad (3.5h) in length, but also contained shorter/longer peptides to provide competition for 3.5h members (see Figure 2). Based on MiSeq deep sequencing data of $>5$ million peptides (of which 99.9 % of were found to be unique), 39 % of peptides were 3.5-heptads in length, with 17 % of peptides longer and 44 % shorter than 3.5h; of the latter, 20 % comprised either single or dinucleotide deletions, or were missing a single codon (see SI). This mixed-length library was designed to include highly diverse sequences while simultaneously varying the length of the peptide to probe the ability of smaller sequences to compete effectively against larger sequences predicted to
form more points of contact with the target. In particular, this library probed whether a peptide shorter than a previously-derived binder, FosW, could be selected. Moreover, the CIS→PCA was employed to deliver a proof of principle of this technique, providing a range of peptide lengths and therefore expected affinities. The theoretical diversity of the library, designed to extensively explore functional space, was $7.3 \times 10^{17}$, and therefore exceeded the maximal screening capacity of CIS display. However, to demonstrate CIS→PCA feasibility, the primary focus was to screen a highly diverse library containing a high probability of cJun binding peptides, such that partial library coverage was satisfactory. Using FosW as a design scaffold, library design options were based on those commonly found in the CC domains of bZIP proteins (Figure 2). Library construction utilised the ProxiMAX codon-by-codon randomisation technique, which avoids peptide-encoding redundancy, bias for selection of overrepresented peptides, and subsequent loss of library diversity. Construction of 3.5h peptides comprising N-terminal truncations of FosW were rationalised on the basis of poor helicity as evidenced by in silico analysis of per residue helical propensity using the Agadir algorithm (Figure S2), and the expectation that a cJun binder would require appreciable helicity in order to be dimerisation competent.

The CIS display library contained residue options to promote and maintain the parallel dimeric CC motif, together with some more unusual options. For example, at core a positions, hydrophobic options (I/L/V) were combined with aromatic residues (F/W/Y; 5% total frequency). These options were mirrored at d positions, which had been fixed as Leu in all other PCA screens against AP-1 to date and randomised here to provide further diversity. These options were hypothesized to provide an increased opportunity to select residue contacts that maximise binding enthalpy and can therefore compensate for diminished length within shorter peptides, provided bulky side-chain packing could be facilitated. The library a2 position was fixed as Asn to provide parallel homotypic dimerisation with the cJun a3 Asn. Charged/polar and atypical hydrophobic residues were offered at solvent-exposed b, c and f positions (E/K/R/Q/L/I/A) to provide opportunities to further promote helicity and solubility as well as further intra-molecular stabilisation via favourable sidechain–sidechain interactions. Finally, e and g positions were given charged/polar (E/K/R/Q) options to promote interhelical interactions as well as solubility and/or helicity. From this
library it was anticipated that a novel peptide sequence with high affinity for cJun would be selected.

**Library Screening Overview**

CIS display \[^{22}\] and PCA \[^{21}\] library display and selection systems were combined for the first time to allow high throughput and thorough screening of an expansive library to identify novel antagonists of cJun. The success of the approach was dependent on relevant selection pressures in the two systems being compatible such that peptides that bind cJun in vitro can also do so under *in cellulo* conditions. CIS display, cloning of CIS display hits into PCA vectors, and subsequent PCA screening of the remaining library (Figure 1) was considered the most accessible and facile strategy in this proof-of-concept trial, to achieve a simple combined system using standard molecular biology and microbiology techniques.

**Library Screening using CIS display**

CIS display screening of $6.1 \times 10^{10}$ peptides meant that the number of 3.5h and 4.5h peptides ($2.4 \times 10^{10}$ and $1.04 \times 10^{10}$) was $\geq 10,000$ fold larger than would be achievable with PCA alone (approx. $1.0 \times 10^6$). Library peptides were subjected to CIS display against an immobilised synthetic cJun CC region peptide, at two different selection stringencies ('low' and 'high' amounts of target presented to the expressed peptide library) to cover a range of possible peptide affinities for the target. CIS display selection efficacy was monitored using deep sequencing of the recovered DNA and resulted in 25,000-72,000 unique sequence reads from each of the two stringencies (recovered after panning rounds 3 and 4). Of these, 85–131 peptides were observed $\geq 50$ times and therefore considered to be sufficiently enriched for further analysis.

**Library Screening using PCA**
The enriched clonal output from the library was subcloned into the PCA system (see SI). This resulted in extensive (≥95%) coverage of remaining peptide diversity, as determined by counting colonies of transformed cells (see SI, Equation 1), to allow any peptide to be selected by PCA regardless of rank within the CIS display enrichment. Initially, PCA consisted of Single-step selection on minimally nutritious M9 medium plates. Following this, sequential passaging of cells in M9 minimal media was performed to provide Competition Selection. This further narrowed the pool of selected peptides to a single sequence by amplifying subtle differences in target affinities between peptides to drive selection of the highest affinity binder \(^{21}\). Both approaches require library–target interaction and concomitant refolding of mDHFR for bacterial growth and peptide selection.

PCA screening that followed low stringency CIS display selection rapidly isolated a single peptide from the ~25,000 library members to enter from the CIS display screen. The peptide, named ‘CPW’ (CIS→PCAWinner), was identified via Sanger sequencing. Reassuringly, this peptide was found to generate the fastest growing colonies at the Single-Step selection phase, indicating the highest recombined mDHFR activity, and was also the monoclonal sequence to emerge after 2-5 rounds of Competition Selection (Figure S3). This sequence emerged as the overall CIS→PCA winner from position 71 in low stringency CIS display (panning round 4 deep sequencing) based on the observed frequency of hits.

Interestingly, for the high stringency CIS display selection, CPW was not identified in the sequenced reads. Rather, the highest-ranking peptide after high stringency CIS display selection also afforded fastest colony growth to win the corresponding high stringency PCA Competition Selection. This peptide (named “CIS1”) was also the highest ranked during low stringency CIS selection, but was ultimately outcompeted by CPW in the following low stringency PCA. During the CIS→PCA process, sequencing data was used to follow selection efficacy (selection round fold-enrichment) and \textit{in silico} prediction algorithms were used to evaluate parameters that describe peptide-cJun binding capabilities (Table S5). Most of the enriched peptides from the low and high stringency CIS display selections did not survive the corresponding PCA, demonstrating the importance of transferring all peptides from CIS.
to PCA. Only 4.5h (32mers + capping motifs = 37mers) library members emerged from PCA screens, supporting the hypothesis that only longer peptides, capable of forming a greater number of specific interactions or that are inherently more stable with cJun, are generally selected.

**CPW sequence analysis**

The residue selections in CPW infer interhelical interactions with cJun residues, as well as intrahelical interactions that stabilise the CPW helix into a parallel dimeric conformation with cJun (Figure 3). Evidence for this comes from the structurally related cJun-FosW structure that we have recently reported (PDB ID 5fv8), as well as the wild-type cJun-cFos (PDB ID: 1fos) interaction, and the fact that PCA selects for parallel dimeric interactions. Peptides adopting higher-order or multiple oligomeric states, or that display the incorrect helix orientation, are highly unlikely since PCA only enriches for peptides that align in a parallel orientation (owing to linker length) during selection and ensures that specific helices are dimeric.

Given the library options provided and where reference data was available, the majority of selected core residues corresponded to those predicted to contribute most to the dimerisation free energy. For example, Ile was selected at three out of five $a$ positions and also, interestingly, three out of five $d$ positions. Also unusually, Phe was selected at position $a4$, where it packs against $a'4$ Ala. Phe has the potential to form van der Waals interactions with Ala; however, the contribution of this pairing to stability has not been studied, nor has to what extent this bulky side chain can be accommodated in a dimeric core.

At $e$ and $g$ positions, three of five positions appear to form interhelical interactions, as expected. Residues selected at $g1$, $e2$, and $g3$ are expected to make interhelical contact, and $e3$ Gln could perhaps make contact with the cJun $g'2$ Gln, with these four positions featuring the most optimal residues for dimerisation free energy. At the remaining $e4$ position, Glu is located opposite Thr, a combination that has not previously been scored but that is unlikely to make a large energetic contribution. Selection of Arg at $g2$ is located opposite Ala and therefore predicted to make only a
small contribution to dimerisation free energy. Arg may, however, participate in a stabilising g–a' interaction with the α3 Asn of cJun, as has been observed in the cJun-cFos crystal structure\(^a\), where it functions to improve dimer stability. Alternatively, Arg may interact favourably with either the solvent or g3 Glu within CPW. This has been observed at the corresponding position within the cJun-FosW crystal structure\(^b\). The Asn-Asn pair that is predicted to aid dimeric specificity is preserved via selection of Asn from six alternative options at α3 in the cJun-CIS1 peptide but lost in cJun-CPW, where Ile was selected.

Possible intrahelical interactions between b, c and f residues could contribute favourably to dimerisation free energy, as well as more generally to CPW solubility and α-helicity, through stabilisation in a helical conformation. An expected preference for Glu, Lys, Gln, Arg and Ala rather than Ile or Leu was observed, with predominance of Gln (seven positions) over Glu (three positions), Lys (two positions) and Ala (one position), and no instances of Arg. The single Ala residue at b4.5 may aid helicity in this poorly helical region of CPW (Figure S2)\(^c\), whilst other selections in these positions require future investigation.

**CPW–cJun in silico prediction**

*In silico* algorithms were used to predict the likelihood of the CPW–cJun interaction, and whether this complex is anticipated to form a CC rather than an alternative binding mode (Table S5). These predictions were compared against those for previously identified cJun-binders, and against CIS1–cJun to identify the highest affinity binder. Predictions by these algorithms are scaled to real, measured interactions between analogous sequences, and thus provide a more representative global view of interaction capability than can be gained from manual comparison of pairwise interaction preferences. Agadir\(^d\) predicted 15 % helicity for CPW, lower than that of FosW (26 %), 4hFosW (22 %) or FosW\(_{\text{CANDI}}\) (22 %), but much higher than native cJun binders (*e.g.* cFos = 3.5 %) and CIS1 (5.4 %), indicating potential for increased affinity for cJun (see also Figure S2). bCIPA\(^e\) was used to predict the interaction stability of CPW–cJun, and calculated a T\(_m\) value of 53 °C. This value indicated promising potential for high affinity (nM) binding, and reassuringly was well above natural AP-1
CC interaction affinities (Tm values -4 °C to 18 °C ). CIS1 on the other hand was predicted to interact less favourably (T_m 44 °C). Finally, Base Optimised Weights (BOW) 40 similarly predicts favourable CPW–cJun interaction relative to known binders and CIS1.

**CPW–cJun Characterisation**

Following solid-phase synthesis and purification, *In vitro* biophysical characterisation of cJun binding was performed using circular dichroism (CD) spectroscopy and isothermal titration calorimetry (ITC) to assess helicity, binding mode and interaction affinity.

**Circular Dichroism Experiments**

CPW demonstrated high affinity binding to cJun *in vitro* to form a CC as monitored by CD using 150 μM total peptide concentration. As confirmed from CD scans (see Figure 4B, Table 1), CPW was found to form a heteromeric complex with cJun, displaying a strong α-helical signature with characteristic minima at 222 and 208 nm. The 222:208 ratio was indicative of formation of a CPW–cJun CC (222:208 ratio >0.9, 53% helicity) 41, 42. Moreover, the heteromeric CD trace far exceeded the non-interaction calculated average of homomeric CPW CCs (slightly more helical than CPW–cJun) and cJun alone (mixed random coil and helical spectrum), which together with ITC data provides firm evidence for the CPW–cJun interaction.

CPW or CPW:cJun mixtures were subjected to thermal denaturation and the CD signal monitored at 222 nm for loss of helical structure. Both provided cooperative thermal denaturation profiles indicating formation of a CC structure (Figure 4), fitting extremely well to a two-state unfolding model 43. The highly helical CPW–cJun CC was thermally stable, with a T_m of 63 °C, displaying a higher T_m, lower K_d and more favourable ΔG than the respective CPW homotypic interaction. The CPW–cJun thermal denaturation profile far exceeded the non-interaction average with regard to helicity as well as the calculated T_m. CPW displayed higher helicity (+22 %) and a much higher T_m (+29 °C) than CIS1, supporting CIS→PCA low stringency selection of CPW over CIS1.
The thermal stability of CPW–cJun compares favourably with CCs of previously derived cJun antagonists (Table 1): the FosW–cJun $T_m$ is identical to that of CPW–cJun, whilst that of 4hFosW–cJun and FosWCANDI–cJun are less stable than CPW–cJun by 14 °C and 11 °C respectively. Compared to the native cFos–cJun AP-1 complex, CPW–cJun is more stable by 47 °C, indicating that inside the cell, CPW would be able to outcompete cFos (or indeed any other Fos/Jun family proteins) to create stable complexes capable of preventing cFos–cJun AP-1 formation.

**Isothermal Titration Calorimetry**

Isothermal titration calorimetry (ITC) experiments were used to more accurately evaluate binding free energy and deconvolute binding into component enthalpic and entropic contributions $^{27, 44, 45}$ (Figure 4 and Table 1). ITC further confirmed the CPW–cJun interaction; titrating cJun into CPW elicited the expected sigmoidal binding curve of a high affinity interaction, with the fit deriving a $K_d$ of 750 ± 270 nM. This translates to a $\Delta G$ of -8.2 ± 0.21 kcal/mol at 20 °C. Enthalpic contributions to $\Delta G$ dominated binding (-7.4 ± 0.67 kcal/mol) relative to the entropic term (0.81 ± 0.81 kcal/mol). These values fit comfortably into the range previously observed for related CC interactions with cJun, including FosW$^{11}$ (45 nM), 4hFosW$^{27}$ (480 nM) and cFos (26.6 μM), and the enthalpically-driven CPW–cJun interaction is typical of CCs $^{27, 46}$. Indeed the predominance of the enthalpic component to CPW–cJun interaction suggests that enthalpy is the dominant driving force for library peptide interaction with cJun, and hence why a 4.5h winner rather than a shorter peptide was selected. CPW displayed a 34-fold lower $K_d$ by ITC than CIS1, again supporting its selection by CIS→PCA. Compared to the successful FosW–cJun interaction, the $\Delta G$ indicates high affinity for cJun suggesting that CPW, like FosW and 4hFosW, would be expected to outcompete cFos for cJun in cellulo to represent an effective antagonist of cFos–cJun AP-1 formation.

**DISCUSSION**
In vitro and in cellulo library screening approaches have been combined for the first time, by using CIS display and PCA, with the aim of assessing whether together the attributes of both systems would constitute an effective system for identification of novel AP-1 antagonists. It was expected that, if compatible, CIS→PCA would benefit from considerable overlap in the capabilities to surmount shortcomings of the isolated respective systems, leading to cJun antagonists with attractive cellular stability isolated from extremely large peptide libraries. A disadvantage of in vitro systems is the likelihood of false positives that can result from non-specific interactions with other components of the assay. This can be partially overcome by addition of more target (‘low’ vs. ‘high’ stringency) and blocking agents, such as BSA. However, this results in lower primary selection stringency, and potentially reduced target affinities of winning peptides. In vitro systems can also potentially lead to enrichment of non-specific (often highly hydrophobic) binders due to the presence of non-target molecules. CIS→PCA allows initial selection of a library under relatively low stringency conditions (CIS display), thereby retaining valuable binders by avoiding requirement for unattainably high affinity, and then subsequently increases selection stringency (PCA) to isolate the highest affinity binders in a cellular environment relevant for the selection of therapeutic characteristics. In fact, PCA inherently selects against non-specific binders as these do not afford rapid host cell growth through sufficient specific interaction with the target and mDHFR recombination. Therefore, it was expected that selection pressures in CIS display and PCA would complement each other to provide the most favourable elements of both systems when combined.

In this study, the ‘low stringency’ CIS selection conditions employed appear to have been more optimal for the affinities of peptides within the screened library. This generated a primary selection stringency that was appropriate to enrich the CPW peptide sufficiently such that it could be transferred to PCA and demonstrate specificity of binding in this assay. Conversely, the ‘high stringency’ CIS condition likely created a primary selection stringency too high for peptide affinities, leading to the inferior CIS1 being selected as the best binder. It is also possible that CPW was not sampled in the ‘high stringency’ CIS condition. The low stringency CIS condition corresponded to cJun at a solution concentration of 290 nM, such that peptides with $K_d \leq 290$ nM would be expected to be selected. Compared to the 750 nM $K_d$ of CPW and 25 μM $K_d$
of CIS1, this suggests that the stringency of this selection could be lowered further in the future to limit potential loss of peptides that could otherwise demonstrate desirable affinity and specificity in PCA. Prior to this study it was unknown whether selection conditions for peptides against a target were sufficiently different during in vitro and in vivo systems to preclude their successful combination. Here we demonstrate for the first time that a cell-free in vitro (CIS display) and in cellulo (PCA) library display system can be effectively combined to derive high affinity cJun binders such as CPW from very large libraries of relatively small peptides.

Construction of the CIS→PCA library resulted in a high diversity library of varying length through the modular construction of DNA corresponding to individual heptads. This meant that longer peptides could be preferentially selected over those with fewer residues by forming more points of interaction with the target (and thus an increased interaction affinity). CIS display enrichment of predominantly 4.5 heptad peptides, and the CIS→PCA selection of CPW, also a 4.5 heptad peptide, supports this. This corroborates the idea that affinity is spread along the length of the cJun interface, making identification of shorter peptides of high affinity more challenging. It is further supported by the dominant role that enthalpy plays in the interaction affinity of CPW–cJun, as with many coiled coils.

The initial library was designed to identify an alternative template to FosW for further peptide development efforts, and to probe the favourability and function of aromatic residues at core positions, while offering options that might aid stability, solubility and helical propensity at poorly understood b, c and f positions. CPW represents such an alternative template, whilst selection of a Phe residue at the a4 position of CPW and selections at b, c and f residues will contribute to our growing knowledge of the structural and sequence determinants of CC formation, which can be exploited in the design of new libraries for selection of modulatory peptides of disease-related CCs. CIS→PCA creates exciting prospects for future assay development and refinement. Combining the two approaches brings additional stringency and the ability to isolate individual sequences via competition selection within cells. This is expected to significantly reduce the number of lead molecules to emerge, with those that do predicted to display improved therapeutic properties. CIS→PCA also has the
considerable potential to provide an accessible and attractive alternative to powerful but complex in vivo phage display (biopanning), where library peptides are selected for affinity and desirable drug-like properties through administration to animal disease models \(^47\). In the future, the concept of CIS→PCA could be translated to combination of an in vitro selection system with yeast two hybrid or other cell-based selection systems more tailored towards expression of eukaryotic proteins of more complicated post-translational modifications. This would further broaden the applications and benefits of CIS→PCA to make it an even more valuable selection strategy.

**CPW as a novel AP-1 antagonist**

The affinity of CPW for cJun is similar to that of FosW and 4hFosW, but novel and non-canonical in sequence, which may yet provide useful design features for future antagonists. The affinity of CPW is lower than that of computationally designed Jun-d1 and peptide-DNA conjugate C2ds (6 nM and 362 nM as determined by solution fluorescence resonance energy transfer and fluorescence anisotropy), and A-Fos, a bZIP CC peptide with an acidic extension (0.03 nM as estimated by CD thermal denaturation), assuming the equivalence of different biophysical techniques. However, the advantage of CPW is its small size; at 37 residues (32mer plus five residues of helix-capping motifs), CPW is shorter than Jun-d1 by five residues, similar in length to 35mer C2ds but lacking 18 bp of conjugated dsDNA, and is shorter than A-Fos by 26 residues. Accordingly, the affinity of CPW for cJun compares well with these molecules when antagonist length is taken into consideration, whilst being of a smaller size that may be more attractive for further therapeutic development. Modifications to improve desired properties while decreasing the molecular weight might include the introduction of helix-inducing constraints at solvent exposed regions to allow further downsizing without loss of intrinsic helicity \(^18\). Finally, compared to T-5224, a small molecule DNA-binding inhibitor discontinued at Phase II clinical trials and demonstrating an in cellulo IC\(_{50}\) of approx. 10 μM \(^48\), the in vitro affinity of CPW is >13,000 greater, making it a promising candidate for in cellulo efficacy at least akin to that demonstrated for T-5224. Further, by targeting the CC domain of cJun rather than the DNA-binding domain – which is highly homologous to other bZIP transcription
factor families $^{49,49,50}$ – CPW could be more specific to cJun-containing oncogenic AP-1 compositions than T-5224.

CONCLUSION

The CPW peptide was selected from a library of $6.1 \times 10^{10}$ peptides, one of the largest libraries screened against cJun to date. In the example presented here, CIS display selection was able to very effectively reduce the library size from $10^{10}$ to approximately $10^5$, making this narrowed pool very accessible to further screening via PCA, which ultimately identified CPW as the highest affinity binder ($K_d$ 750 nM) with \textit{in cellulo} activity. In comparison, the best peptide candidate from CIS display alone (CIS1 - highest ranked peptide in high and low stringency CIS display) displayed a 34-fold weaker $K_d$ for cJun. The superior binding of CPW relative to top-ranked peptides selected by CIS display alone thus supports the original hypothesis that CIS→PCA combination may represent a more attractive system than CIS display or PCA alone. Moreover, the results described here suggest that inside the cell CPW can outcompete cFos for cJun to antagonise cFos–cJun formation and AP-1 oncogenic activity. Further investigation and development of this peptide, as well as refinement of CIS→PCA, could generate valuable antagonists for the future therapy of cancers featuring AP-1 dysregulation, or aid in the development of other modulatory peptides towards this goal.

METHODS

Library construction by ProxiMAX $^{28}$, CIS display $^{22}$ and PCA $^{21}$ were performed as previously described, with minor modifications detailed in the SI. CD spectroscopy was performed on the winning peptide alone or a 1:1 stoichiometric mixture of peptide:cJun at 150 μM total protein concentration (P$_t$) in potassium phosphate without potassium fluoride (“low salt”) buffer. Raw ellipticity was converted to mean residue ellipticity (MRE; deg cm$^2$/dmol) to normalise for different peptide lengths and concentrations as described previously (see SI). ITC was performed by injecting 620 μM cJun into 70 μM winning peptide again in “low salt” buffer. Change in heating power
(μcal/sec) with each injection was integrated to generate binding isotherms, from which ΔH, ΔS, ΔG and $K_d$ values were derived by non-linear least-squares fitting as described previously using Origin software (OriginLab, MA, USA) (see SI).

**AUTHOR CONTRIBUTIONS**

D.B. conducted experiments, and synthesized, purified and characterised peptides and Jun. J.M. and C.U. directed the research. All authors participated in experimental design, analysis of the data, and writing the paper.

**COMPETING FINANCIAL INTERESTS**

CU and LF were previously employed at Isogenica, Essex, U.K. Other authors declare no financial or commercial conflict.

**ACKNOWLEDGEMENTS**

J.Mason and D.Baxter thank the BBSRC for a Case Studentship (BB/J011290/2). J.Mason is also grateful to Cancer Research UK for a Career Establishment Award (A11738).

**SUPPORTING INFORMATION**

Information is provided on Materials (Table S1), Oligos (Table S2-3), ProxiMAX-CIS→PCA overview, Ligations, PCR amplifications, ProxiMAX Library Construction PCRs, CIS display selection ‘recovery’ PCRs, Deep Sequencing Sample Preparation PCRs, CIS→PCA cloning PCRs, Restriction digests, dephosphorylation, Plasmid extraction,
DNA purification and desalting, DNA quantification, deep sequencing, and Sanger sequencing (Figure S3), and RP-HPLC and MS data (Figure S4). For library screening additional information is provided on CIS display, microbiology media (Table S4), bacterial transformation, and PCA. In silico analysis information on Agadir (Figure S2), bCIPA and Base Optimised Weights is also provided.

REFERENCES


**Figure 1:** CIS→PCA. Library construction and selection of cJun binding peptides by sequential CIS display and Protein-fragment Complementation Assay (PCA) 21, 22, 28. Black arrows indicate the path of progression through CIS→PCA. CIS display and PCA schematics adapted from sources. For further details of ProxiMAX, CIS display and PCA, see SI and Figure S1.

**Figure 2:** Peptide library design. The mixed length library was built in a modular fashion from DNA corresponding to Heptad 1, 2 and 4.5 cassettes (each is underlined) constructed using the ProxiMAX codon-ligation approach 28. Within each cassette, codons corresponding to amino acids were incorporated at equal frequencies, except aromatic residues F, W and Y, which were introduced at a total frequency of 5 % (split equally). Heptad positions are in bold italics, library theoretical diversity is in bold text. Note that for internal heptads, 1-3 copies of the Heptad 2 library module were use to create the highly diverse mixed-length peptide library.

**Figure 3:** Possible interactions stabilising a dimeric CC interaction between CPW and cJun. A DrawCoil helical wheel diagram 51. Core interhelical $a-a'$ and $d-d'$ interactions are shown as black arrows, and expected stabilising interhelical $e_i-g_{i+5}$ and $g_i-e'_{i+5}$ interactions are shown as blue dashed lines. A possible interhelical polar-polar interaction ($e_3-g'_2$ Q–Q) is shown as an orange dashed line. All possible $i\rightarrow i+3$ and $i\rightarrow i+4$ intrahelical interactions between outerface $b$, $c$ and $f$ residues (some mutually
exclusive) are displayed with orange dashed lines. Possible intrahelical interactions between outerface residues and \textbf{e} and \textbf{g} residues are not shown. Residues are coloured grey for hydrophobic, red for positively charged, orange for polar, and blue for negatively charged. **B** Linear sequence view of helical wheel diagram in **A**. Heptad register shown in italics. **C** Contributions to interaction $\Delta G$ from selected residues in CPW assuming interaction with cJun residues, with Ala–Ala as the reference state (mimicking no selection). $\Delta \Delta G$ energies from Acharya \textit{et al.} \textsuperscript{36} and Krylov \textit{et al.} \textsuperscript{38}, with two values depending on residue orientation. ND = not determined.

**Figure 4:** CIS$\rightarrow$PCA selected CPW binds cJun with high affinity, and higher affinity than CIS display selected CIS1. Biophysical characterisation of CPW or CIS1 alone and with cJun by CD spectroscopy scans (**A** and **D**) and thermal denaturation (**B** and **E**), and ITC (**C** and **F**). CD data is reported as change in mean residue ellipticity (MRE; units deg cm$^2$ dmol$^{-1}$, to allow for comparison between peptides of different lengths), as a function of CD ellipticity over the wavelength range 200 – 300 nm or at 222 nm with temperature. Peptides interact with cJun where MRE and/or melting temperature ($T_m$) for peptide:cJun mixes exceed that of the calculated average trace of peptide and cJun only traces (which represents non-interaction). ITC raw isotherms (top panels) and fitted data (bottom panels) (both baseline corrected). On the fitted data plot, the solid line represents the fit generated by non-linear least-squares fitting. See \textsuperscript{19} for equations using for fitting.
Table 1: CPW displays high affinity binding to cJun, similar to that of previous cJun antagonists, substantially higher than native cFos, and higher than CIS display selected CIS1.

<table>
<thead>
<tr>
<th>Complex</th>
<th>$T_m$ (°C) $(CD)$</th>
<th>Peptide–cJun Fractional Helicity (%) (CD)</th>
<th>$K_d(20°C)$ (nM) (CD)</th>
<th>$\Delta G(20°C)$ (kcal/mol) (CD)</th>
<th>$K_d(20°C)$ (nM) (ITC)</th>
<th>$\Delta G(20°C)$ (kcal/mol) (ITC)</th>
<th>N interaction stoichiometry (ITC)</th>
<th>$\Delta H(20°C)$ (kcal/mol) (ITC)</th>
<th>$\Delta S(20°C)$ (kcal/mol) (ITC)</th>
<th>Peptide–cJun $\theta_{222/208}$ (20°C) (CD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cFos–cJun</td>
<td>16</td>
<td>28</td>
<td>320,000</td>
<td>-5.5</td>
<td>27,000</td>
<td>-6.1 ± 0.39</td>
<td>-0.82 ± 0.36</td>
<td>5.3 ± 0.53</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>FosW–cJun</td>
<td>63</td>
<td>37</td>
<td>4.0</td>
<td>-11</td>
<td>39 ± 11</td>
<td>0.99 ± 0.08</td>
<td>-9.9 ± 0.16</td>
<td>-10 ± 0.42</td>
<td>-0.46 ± 0.46</td>
<td>1.0</td>
</tr>
<tr>
<td>4hFosW–cJun</td>
<td>49</td>
<td>60</td>
<td>490</td>
<td>-8.5</td>
<td>480</td>
<td>1.1 ± 0.01</td>
<td>-8.8 ± 0.10</td>
<td>-14 ± 0.20</td>
<td>-5.3 ± 0.20</td>
<td>0.99</td>
</tr>
<tr>
<td>FosW$_{CANDI}$–cJun</td>
<td>52</td>
<td>54</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>1.0</td>
</tr>
<tr>
<td>CPW–cJun*</td>
<td>63</td>
<td>53</td>
<td>0.27$^*$</td>
<td>-13</td>
<td>750$^* \pm$ 270</td>
<td>0.48 ± 0.03</td>
<td>-8.2 ± 0.21</td>
<td>-7.4 ± 0.67</td>
<td>0.81 ± 0.81</td>
<td>0.95</td>
</tr>
<tr>
<td>CIS1–cJun*</td>
<td>34</td>
<td>31</td>
<td>14,000</td>
<td>-6.5</td>
<td>25,000 ± 13,000</td>
<td>0.11 ± 0.02</td>
<td>-6.2 ± 0.29</td>
<td>-15 ± 4.42</td>
<td>-9.2 ± 4.6</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Thermodynamic parameters for CPW in interaction with cJun compared to cJun antagonists FosW$^{46}$, 4hFosW$^{37}$, and FosW$_{CANDI}$, and native cFos. CPW is also compared to less successful CIS→PCA high stringency winner CIS1, the most frequent peptide in CIS high and low stringency selection. Data from CD and ITC measurements. θ is raw CD ellipticity (mdeg). Fractional helicity was calculated as previously$^{19}$. $\Delta S$ is calculated as $\Delta H – \Delta G$ from the Gibbs-Helmholtz equation (see$^{19}$). CD values are from representative single measurements, typically reproducible in biological replicates to ± 1 °C for $T_m$, within 5% for fractional helicity and 222:208 ratio (θ222/208), and within 10% for $K_d$ and $\Delta G$ (data not shown). ITC values are the arithmetic mean of two to three independent titrations ± SDs, except values for 4hFosW–cJun, FosW$_{CANDI}$–cJun and cFos–cJun (single titrations and fitting errors). *'low salt' buffer (10 mM KH$_2$PO$_4$/K$_2$HPO$_4$, no KF), other peptides assayed in KPP buffer (*low salt*+ 100 mM KF). "ND" = not determined. Values given to 2 s.f. $^*$Indicates a discrepancy between CD and ITC derived $K_d$ values; for all comparisons the more accurate ITC derived $K_d$ has therefore been adopted.