Motifs for making tricycles

Christian Heinis[1]

[1]Institute of Chemical Sciences and Engineering, School of Basic Sciences, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

Cyclic peptides are a rising molecular format for developing ligands to proteins of interest which can be used as research tools (such as binders for chemical probes) and for drug development. Ligands based on cyclic peptides can be efficiently identified by screening billions of random sequences generated using techniques like phage or mRNA display. Any binders that are identified can then be conveniently produced by chemical synthesis. Cyclic peptides can bind to a target with higher affinities compared to linear ones due to a restricted conformational flexibility, which reduces the entropic penalty upon target binding. The effect of cyclization on affinity is stronger the more the peptides are constrained by intra-molecular bridges. This trend can be seen by comparing linear, monocyclic, bicyclic, and tricyclic peptides (Figure 1a). Thus, peptides cyclized with three or more chemical bridges are of great interest for generating binders to challenging targets.

Tricyclic peptide ligands also occur in nature such as cystine knots[1] and some have been exploited as scaffolds for engineering binders to new targets, with remarkable recent successes[2]. However, the applied libraries were based on single tricyclic peptide scaffolds. Developing libraries based on multiple, randomly generated tricyclic peptide folds would enable screening much greater structural diversities — and so potentially find better binders — but so far this has been difficult due to technical challenges in forming multiple site-specific connections in random peptides.

Writing in the *Journal of the American Chemical Society*[3], a team led by Chuanliu Wu have now established a simple and robust strategy for efficiently generating and screening large libraries of structurally diverse tricyclic peptides. This approach yielded binders with remarkably good affinities to several proteins considered challenging for ligand development. The strategy is based on phage display of linear peptides containing six cysteines that are oxidized to form three disulfide bridges and thus tricyclic peptides. A major challenge with this strategy was that six cysteines in one peptide can pair in 15 different ways and that the complex mixture of products could hinder an efficient enrichment of binders as well as complicate subsequent peptide synthesis. Chuanliu Wu and colleagues have elegantly solved this problem using disulfide-directing peptide motifs that allow the formation of preferred disulfide pairs to produce from each linear peptide one tricyclic peptide as main product.

Peptide motifs bearing twin cysteines such as CPPC and CXC (where X can be varied) have a propensity to form dimers with a pair of inter-motif disulfide bonds instead of forming intra-motif bridges. Wu and colleagues had previously taken advantage of such motifs to control the disulfide pairing in peptide libraries[4,5]; however, the small number of variable amino acids within the motifs limited the structural diversity of subsequent peptide libraries. In their new work, the team discovered a disulfide-directing peptide motif with the sequence CPXXC, which preferentially pairs with a second such motif in anti-parallel orientation and in which two positions (X) can be varied. The new motif was inspired by human plasma fibronectin that forms an antiparallel dimer via the short peptide CPIEC. Incorporating two such fibronectin-based motifs into peptides favors the formation of one of the three possible disulfide pairings (Figure 1b). Substitution of the three middle amino acids in the motif CPIEC showed that mainly proline was important for directing the disulfide pairing (CPXXC), which enabled varying two amino acids in peptide libraries.

To generate libraries of tricyclic peptides, Wu and co-workers added two additional cysteines in the loop that connected the two CPXXC motifs (Figure 1c). HPLC analysis of several peptides with this format showed that it led to the formation of mainly one out

of the 15 possible cysteine connections, though some side products appeared with varying frequencies. Despite the side products, the formation of a dominant product meant there was a good chance that phage display selections could work. Thus, a phage display library comprising 440 million tricyclic peptides was prepared, with amino acids randomized in all five of the cysteine-spaced segments (Figure 1c). Panning the library against ephrin type-A receptor 2 (EphA2), fibroblast growth factor receptor 1 (FGFR1), human epidermal growth factor receptor 2 (HER2), and HER3 yielded single-digit to low double-digit nanomolar binders for all targets. The binding affinities achieved are remarkably good, especially considering that the four cell-surface proteins are rather challenging targets.

A key strength of the new approach is that it generates tricyclic peptide libraries that contain a large number of different folds, as opposed to only one 3-dimensional architecture. This contrasts with previous approaches comprised of nature-derived peptide structures that thus generated tricyclic peptide libraries based on single scaffolds. In addition, the new approach diversifies the included amino acids in five segments (spaced by Cys or Cys-Pro), which is much more than in libraries based on natural scaffolds that have regions that cannot be altered as they are needed to direct the folding and disulfide pairing. Instead, here it is the disulfide-directing motifs that automatically guide the structure, so specific amino acid sequences are not required to direct the folding.

Finally, what really makes the new approach highly attractive is its simplicity and robustness. The cyclization occurs automatically when the phage particles are released from the bacterial cells into an oxidative environment, so no care needs to be taken here. This contrasts sharply with methods such as the phage-display selection of linker-cyclized peptides, which does require the careful control of chemical cyclization conditions. It is likely that this simplicity and robustness, combined with the prospect of discovering good binders to difficult targets, will lead to a broad application of the new method for generating peptide-based ligands and potentially developing new classes of therapeutics.

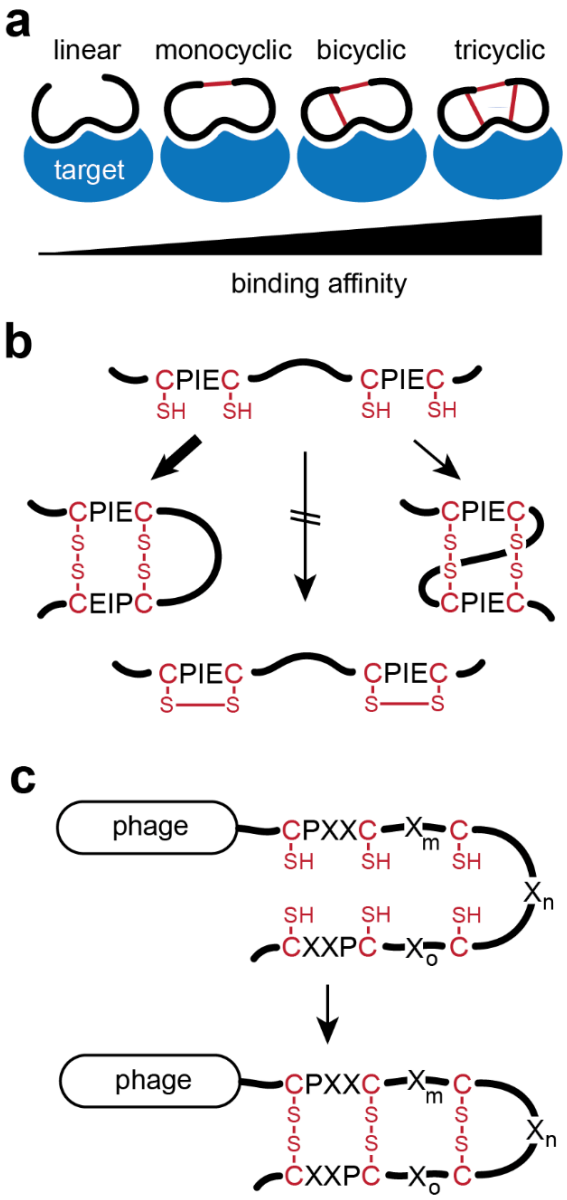

Figure 1 Phage display selection of tricyclic peptides using a disulfide-directing motif. **a**, Cyclization linkers (red) reduce the flexibility of peptides for a smaller entropic penalty upon target (blue) engagement, allowing for higher affinity binding. **b**, The motif "CPIEC", derived from human plasma fibronectin, directs the disulfide pairing and reduces the number of products. **c**, The disulfide-directing motif "CPXXC" is used to display tricyclic peptides on phage. The motif leads to formation of one of 15 possible peptide configurations as main product, which enables the phage display selection of tricyclic peptides to challenging targets.

References

1. Daly, N. L. & Craik, D. J. Bioactive cystine knot proteins. *Curr. Opin. Chem. Biol.* **15**, 362–368 (2011).
2. Hansen, S. *et al.* Directed evolution identifies high-affinity cystine-knot peptide agonists and antagonists of Wnt/β-catenin signaling. *Proc. Natl. Acad. Sci.* **119**, (2022).
3. Lu, S. *et al.* Disulfide-Directed Multicyclic Peptide Libraries for the Discovery of Peptide Ligands and Drugs. *J. Am. Chem. Soc.* **145**, 1964–1972 (2023).
4. Wu, C., Leroux, J.-C. & Gauthier, M. A. Twin disulfides for orthogonal disulfide pairing and the directed folding of multicyclic peptides. *Nat. Chem.* **4**, 1044–1049 (2012).
5. Lu, S. *et al.* Directed Disulfide Pairing and Folding of Peptides for the De Novo Development of Multicyclic Peptide Libraries. *J. Am. Chem. Soc.* **142**, 16285–16291 (2020).