DRUG DISCOVERY

Identifying rescuers of misfolding

High-throughput screening of large libraries of cyclic peptides expressed in bacteria yields rescuers of the pathogenic misfolding of proteins associated with neurodegenerative diseases.

Tobias Langenberg, Joost Schymkowitz and Frederic Rousseau

Many diseases arise as a consequence of the misfolding of proteins. Such pathologies include type-II diabetes, cystic fibrosis, systemic amyloidosis, cataracts, and neurodegenerative diseases such as Alzheimer disease (AD), Parkinson's disease and amyotrophic lateral sclerosis (ALS, a fatal neurodegenerative disease characterized by the destruction of peripheral motor neurons). Protein misfolding leads to loss of protein function (and subsequent degradation), and misfolded species typically aggregate into amyloid-like oligomers. Amyloid aggregates can slowly accumulate in tissues, leading to toxic effects. Because of the worldwide ageing of the population, the social and medical impact of many of these diseases keeps increasing. Nevertheless, despite some progress in therapy development¹, effective treatments remain a significant challenge, especially for neurodegenerative diseases. In AD, as in many other neurodegenerative diseases, the onset of disease occurs well before the detection of the first symptoms, making it especially difficult to tackle disease progression; consequently, treatment is largely restricted to the alleviation of symptoms, which highlights the need for effective disease-modifying therapies².

The toxicity of amyloid aggregates may be countered by different approaches, for example by shifting the ratio between folding and misfolding, by blocking the progression of the aggregation pathway, or by promoting clearance of the aggregates. Reporting in *Nature Biomedical Engineering*, Georgios Skretas and colleagues show that the combination of large libraries of short cyclic peptides produced in bacteria and a simple fluorescence-based genetic assay to measure soluble (and thus not aggregated) proteins fused with green fluorescent protein (GFP) allows for the identification of peptides that increase the fraction of the folded forms of these fusion proteins³. The peptides rescuing the proteins from their misfolded states can constitute the basis of potential therapies for their respective associated diseases. In the authors' case, they used fusion proteins of either amyloid beta (A β) peptide or superoxide dismutase one (SOD1). A β accumulation is a predominant feature of AD, whereas SOD1 misfolding and aggregation is implicated in a subset of patients suffering from ALS.

To generate cyclic tetra-, penta- or hexapeptides in *E. coli*, Skretas and co-authors employed a technique (SICLOPPS; split intein-mediated circular ligation of peptides and proteins) that makes use of self-excising protein domains called inteins that, when properly arranged, have the ability to excise and ligate an intervening peptide⁴. By using degenerate codons, large vector libraries (theoretically, larger than 10 million sequences) can be constructed, and production in *E. coli* ensures that the full library can be screened. The authors used the library to identify peptides that reduced aggregation of the A β -GFP fusion reporter, by analysing via fluorescence-activated cell sorting a population of bacterial cells overexpressing simultaneously the peptide library and the fusion protein. They selected ten random clones with a strong increase in GFP fluorescence, and observed a strong preference for pentapeptides with a specific sequence motif. Deep sequencing of more than 5 million clones with enhanced GFP fluorescence revealed a predominant cyclo-TXXXR motif in this population. Two peptides were then taken forward for functional characterization, one falling into the TXXXR class (TAFDR), and one other (SASPT) unrelated. The authors found that these peptides have the ability to interfere with A β fibril formation in vitro (diminishing the neurotoxicity of both A β 40 and A β 42 in primary mouse hippocampal neurons) and to suppress the toxicity of A β in vivo. The latter observation came from experiments in various *C. elegans* AD models, such as a strain expressing human A β 42 in body-wall muscle cells that develops paralysis at the onset of adulthood. When the authors fed these worms with bacteria overexpressing either of the two cyclic peptides throughout their lifetime, they observed a significant delay in the onset of paralysis.

To demonstrate the general applicability of their approach, Skretas and colleagues applied the same assay to identify inhibitors of SOD1 misfolding, and more specifically to a heavily aggregation-prone mutant of SOD1, A4V. The accumulation of SOD1-A4V is causative in a very aggressive form of ALS (ref. ⁵). As in the case of the peptides targeting A β , deep sequencing revealed a predominant motif, in this case cyclo-TSXXW. The peptide TDSDW was selected for further characterization, and was found to inhibit the aggregation of SOD1-A4V in vitro: the fraction of large aggregated species was much reduced, and the decrease in staining intensity with the conformation-sensitive dye SYPRO orange supports a mechanism of action by which the peptide blocks exposed hydrophobic patches on SOD1.

For the A β -targeting peptides, Skretas and co-authors used docking and molecular dynamics simulations to identify potential binding sites, and concluded that the peptides inhibit aggregation by interacting with different faces of the A β fibre. The

simulations are in agreement with structure–activity analyses showing additional charged residues being strongly disfavoured in the predominant cyclo-TXXXR motif.

The work by Skretas and colleagues is an example of the power of unbiased, high-throughput approaches that are also simple and robust. In neurodegenerative diseases, and in amyloid diseases in general, structural information on (early) toxic aggregates is scarce or non-existent, making modelling-based approaches difficult or impossible. By contrast, approaches such as that of Skretas and colleagues do not require any structural, or even mechanistic, knowledge, but are based merely on an intracellular assay for the desired effect (in this case an increase in the level of GFP fluorescence). Moreover, because aggregating peptides and proteins associated with neurodegenerative diseases are generally difficult to handle, the intracellular nature of the SICLOPPS assay makes it uniquely suited for the identification of potential drugs, and it has been ported to eukaryotic cells for screening in the relevant cellular environment^{6–8}. Compared with the conventional phage-display or mRNA-display methods of (cyclic) peptide selection, the number of sequences assayed by SICLOPPS is smaller, and peptide affinity tends to be lower (possibly owing to the shorter peptides typically generated)⁹. However, longer peptides may not necessarily make for better drug candidates, and even the enormous library size of up to 10¹⁵ candidates in mRNA display libraries is not sufficient to cover the full sequence space of peptides with all 20 amino acids per position at 13-mers (8.19 x 10¹⁶ peptides). Solutions to this problem could be based on iterative screening rounds or on limited design, and are also possible by using SICLOPPS with longer peptides.

Interestingly, even though there was a strong enrichment of specific sequence motifs in the candidate hits, other randomly selected peptides occurring with low frequency were also found to be effective, which means that the strongest sequence motif is not necessarily the one that should be pursued. Additional challenges in the development of protein misfolding and aggregation therapies remain, such as the critical necessity for compounds against neurodegenerative diseases to cross the blood–brain barrier. In addition, since neurodegenerative diseases such as AD are progressive and develop long before overt symptoms appear, anyone at risk may need to take an inhibitory drug on a daily basis from the 5th or 6th decade of their lives, which would require that these compounds have good oral bioavailability and very low toxicity. The emergence of screening methods such as that presented by Skretas and colleagues cannot, of course, solve all these issues at once, but nevertheless provide exciting opportunities for the discovery of new therapeutic molecules for these diseases, and for the development of strategies that stabilize folded protein conformations or that inhibit the kinetics of aggregation of misfolded proteins.

Tobias Langenberg, Joost Schymkowitz and Frederic Rousseau are at the VIB Switch Laboratory; Department of Cellular and Molecular Medicine, KU Leuven, Herestraat 49, PO 802, B-3000 Leuven, Belgium.
e-mail: frederic.rousseau@vib.kuleuven.be

References

1. Gomes, C. M. Protein misfolding in disease and small molecule therapies. *Curr Top Med Chem* **12**, 2460–2469 (2012).
2. MD, P. P. S. *et al.* Seminar Alzheimer's disease. *The Lancet* **388**, 505–517 (2016).
3. Matis, I. *et al.* *Nat. Biomed. Eng.* **1**, XXX–YYY (2017).
4. Scott, C. P., Abel-Santos, E., Wall, M., Wahnou, D. C. & Benkovic, S. J. Production of cyclic peptides and proteins in vivo. *Proc Natl Acad Sci USA* **96**, 13638–13643 (1999).
5. Kaur, S. J., McKeown, S. R. & Rashid, S. Mutant SOD1 mediated pathogenesis of Amyotrophic Lateral Sclerosis. *Gene* **577**, 109–118 (2016).
6. Kinsella, T. M. *et al.* Retrovirally Delivered Random Cyclic Peptide Libraries Yield Inhibitors of Interleukin-4 Signaling in Human B Cells. *Journal of Biological Chemistry* **277**, 37512–37518 (2002).
7. Mistry, I. N. & Tavassoli, A. Reprogramming the Transcriptional Response to Hypoxia with a Chromosomally Encoded Cyclic Peptide HIF-1 Inhibitor. *ACS Synth. Biol.* **6**, 518–527 (2016).
8. Kritzer, J. A. *et al.* Rapid selection of cyclic peptides that reduce α -synuclein toxicity in yeast and animal models. *Nat Meth* **5**, 655–663 (2009).
9. Tavassoli, A. SICLOPPS cyclic peptide libraries in drug discovery. *Current Opinion in Chemical Biology* **38**, 30–35 (2017).

Figure 1 | High-throughput genetic screening for the discovery of macrocyclic rescuers of pathogenic A β misfolding and aggregation. (a) Schematic representation of the utilized bacterial system for the discovery of cyclic oligopeptide rescuers of MisP misfolding and aggregation. pMisP-GFP: plasmid encoding a MisP-GFP fusion; pSICLOPPS-NuX1X2X3-X5: vector library encoding the combinatorial oligopeptide library cyclo-NuX1X2X3-X5. Nu: Cys (C), Ser (S), or Thr (T); X: any of the twenty natural amino acids; NNS: randomized codons, where N=A, T, C or G and S=G or C; P: sorting gate. Reproduced from ref. 3.