Minireview

Patrick Ernst and Andreas Plückthun*

Advances in the design and engineering of peptide-binding repeat proteins

DOI 10.1515/hsz-2016-0233
Received June 12, 2016; accepted August 23, 2016; previously published online September 15, 2016

Abstract: The specific recognition of peptides, which we define to include unstructured regions or denatured forms of proteins, is an intrinsic part of a multitude of biochemical assays and procedures. Many cellular interactions are also based on this principle as well. While it would be highly desirable to have a stockpile of sequence-specific binders for essentially any sequence, a *de novo* selection of individual binders against every possible target peptide sequence would be rather difficult to reduce to practice. Modular peptide binders could overcome this problem, as preselected and/or predesigned modules could be reused for the generation of new binders and thereby revolutionize the generation of binding proteins. This minireview summarizes advances in the development of peptide binders and possible scaffolds for their design.

Keywords: armadillo repeat; directed evolution; protein design; protein engineering; repeat proteins; tetratricopeptide repeat.

Introduction

The specific recognition of proteins and their interaction partners is of great interest in both research and diagnostics, with antibodies being the most widely used detection agents. Within the last years it became clear, however, that many conventional monoclonal antibodies widely used in research have two fundamental problems: first, they are poorly characterized and second, they are almost never molecularly defined by their sequence, as they are not recombinant (Bradbury and Plückthun, 2015). The use of poorly characterized and ill-defined antibodies can not only lead to false or non-reproducible results but also to a great waste of money, which is why researchers are calling out for a solution to the so-called 'reproducibility-crisis' (Bradbury and Plückthun, 2015). We should stress that therapeutic antibodies do not have this problem, as FDA and EMA approval requires both molecular definition and very high quality standards.

One way to address this issue is the use of well-characterized recombinant binding reagents. As this makes the reagent completely independent from immunized animals, there is also no need to stay with the antibody framework, but we can search for alternative binding scaffolds, some of which may have very favorable biophysical properties. Protein design has led to the emergence of numerous alternatives to classical monoclonal antibodies, which can be selected to specifically bind new targets – typically folded proteins (Jost and Plückthun, 2014). But what about peptides as targets?

We define 'peptides' to include unstructured regions or denatured forms of proteins, and this binding thus describes a particular mode of interaction between proteins, in which one partner is not constrained by a three-dimensional structure, and may also give access to interactions with its backbone. This type of interaction is found in many protein-protein interactions. Moreover, a multitude of biochemical assays rely on this, such as protein detection in Western blots, immunohistochemistry, or enrichment of peptides for mass spectrometry.

The interactions of proteins with peptides or unstructured regions of proteins are presumed to be involved in up to 40% of all interactions within the cell (Petsalaki and Russell, 2008). Since the surface of folded proteins is irregular, there is no way around selecting a new set of binders for every folded domain. Unstructured peptides are seemingly simpler structures, and it would be appealing to have at our disposal a set of directly available binding proteins against each peptide sequence. However, it becomes immediately obvious that the number of sequence possibilities for even a hexapeptide ($20^6 = 6 \times 10^7$) makes the logistics of such a proposal look rather unrealistic, and for longer peptides outright absurd. Modular peptide binders would have the potential to circumvent this obstacle. If the

***Corresponding author: Andreas Plückthun,** Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland, e-mail: plueckthun@bioc.uzh.ch
Patrick Ernst: Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

modularity of the binding protein parallels the modularity of the peptide, modules for side chains might be generated, and far fewer of them would be needed. A preselected set of binding modules could then be reassembled at the gene level to bind, in principle, every peptide target sequence (see Figure 1). As recently reviewed (Reichen et al., 2014a), repeat proteins are highly suited to fulfill this purpose, because they can be varied in the number of internal repeats they contain, all carrying positions of variable amino acids on their surface. These can serve as binding sites, and indeed many of the repeat protein domains are involved in binding and signaling in biological systems.

In recent years, more and more examples of artificial binding proteins based on repeat proteins have been reported, both for recognizing folded proteins and for peptides. An extensive overview on the designs and their development has been given by Reichen et al. (2014a). Among the repeat proteins, α-solenoid proteins have been most widely investigated, and examples include ankyrin repeat proteins, HEAT repeat proteins, armadillo repeat proteins (ArmRP) and tetratricopeptide repeat proteins (TPR). Ankyrins served as a template to create designed ankyrin repeat proteins (DARPins) (Plückthun, 2015), while HEAT proteins are the basis of αReps (Urvoas et al., 2010), the latter two classes mostly to derive binders to folded proteins. ArmRPs and TPRs have been used to design peptide binders, as this is also one of their functions *in vivo*. Closed architectures have also been investigated (Orlicky et al., 2003) but we will not consider them further here.

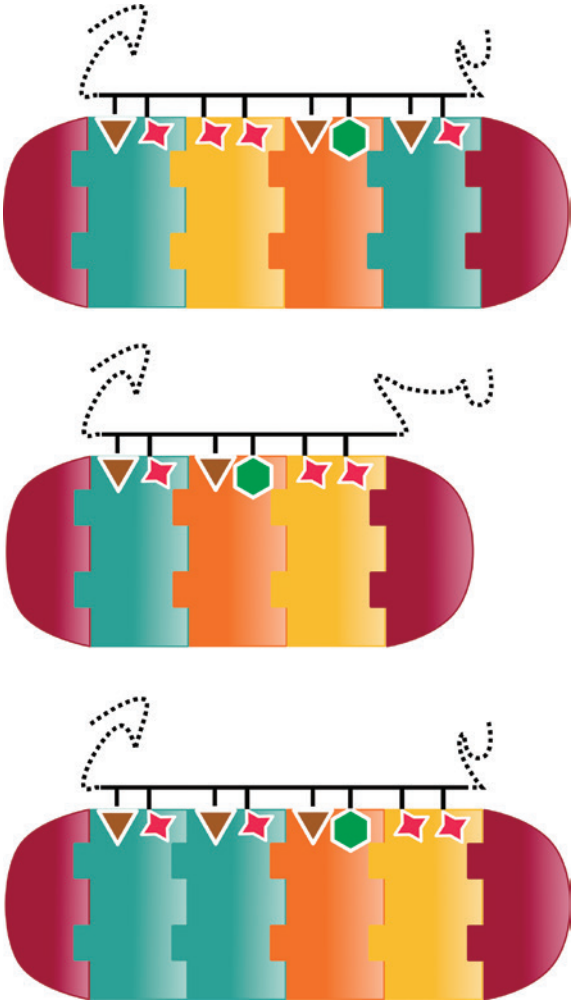

Figure 1: Modular peptide binding.
Principle of modular peptide binding with internal binding modules of different number, arrangement and specificity, binding to unstructured regions of proteins. The hydrophobic core of the binder is shielded by special capping repeats (violet) at the N- and C-terminus, respectively.

In this mini-review we will focus on recent advances in the development of new peptide binders based on ArmRPs or TPRs and their possible applications. Furthermore, we will provide an outlook of possible new repeat protein scaffolds that have emerged due to the development of better computational protein design techniques, and we discuss their potential for the future.

Tetratricopeptide repeat proteins

TPRs consist of repeats of 34 amino acids in length and have been used as a template to design highly stable consensus TPRs (CTPRs) which were able to bind peptide ligands with different sequences (Jackrel et al., 2009; Cortajarena et al., 2010). Furthermore, a detailed biophysical investigation on the scaffold has been carried out to understand the folding and stability of designed CTPRs (see, e.g. Cohen et al., 2015; Millership et al., 2016). To describe the ability for binding to different peptides we will focus here on the design of new pockets and their possible applications.

A binding module of a CTPR consists of three internal repeats and a capping helix at the C-terminus, with binding pockets recognizing the side chains of the pentapeptide MEEVD (Figure 2A; Cortajarena et al., 2010). These designed CTPRs originate from TPR2A, which binds to the C-terminal part of Hsp90 with an affinity of 11 μM (Scheufler et al., 2000). The C-terminal carboxylate group and the aspartate side chain form a dicarboxylate-clamp, a key feature for binding (Scheufler et al., 2000; Cortajarena et al., 2004). This binding pocket was redesigned to recognize a phenylalanine instead of the conserved aspartate (Jackrel et al., 2009). These

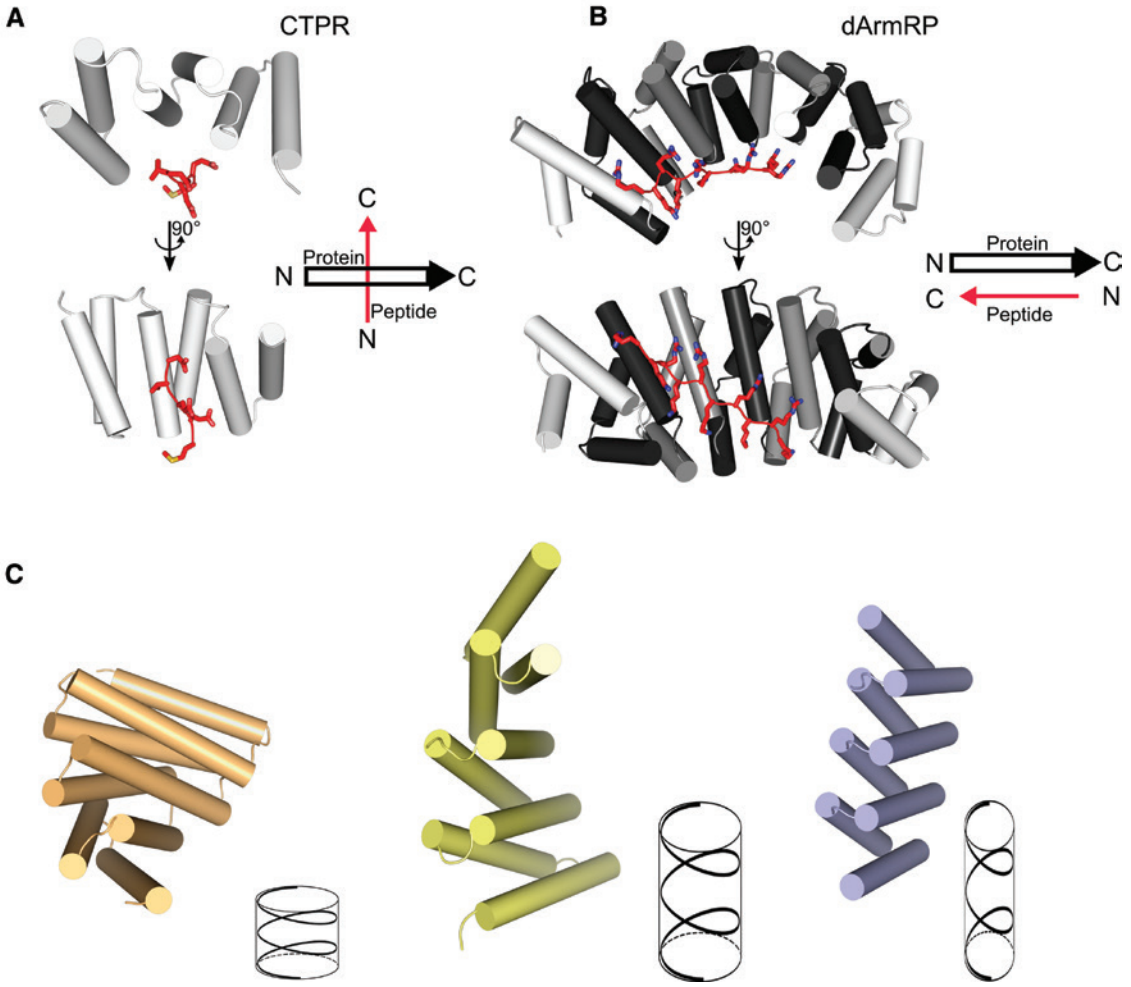

Figure 2: Overview of existing peptide binders and possible new scaffolds based on repeat proteins.
(A) Designed TPR protein CTPR390 (white), binding to peptide MEEVD (red), top- and side-view (PDB ID: 3KD7 chains A and H). The perpendicular binding of the peptide to the CTPR is schematically drawn on the right, with N and C indicating the respective termini of the protein and peptide. (B) Modular ArmRP with five internal repeats (white) binding a $(KR)_5$ peptide (red) with one KR module per internal repeat. Internal repeats are alternatingly colored in black and gray (PDB ID: 5AEI chains A and D). The antiparallel binding of the peptide to the dArmRP is schematically drawn on the right, with N and C indicating the respective termini of the protein and peptide. (C) Examples for new backbone geometries of designed repeat proteins (PDB IDs: 5CWB, 5CWK, 5CWH). A sketch of the different curvatures of the scaffolds, with different rise and angles between the repeats, are shown next to them.

engineered CTPRs have been called tetratricopeptide repeat affinity protein (TRAPs).

Recent developments have shown the ability of redesigned TRAPs to bind to various peptides with micromolar affinities. Starting from a binder recognizing the original sequence MEEVD, a specificity was developed for a phospho-serine residue (Sawyer et al., 2014). The central glutamate pocket was mutated, introducing arginine and lysine, thereby forming a charge complementarity for the negatively charged phospho-serine. This led to some selectivity, as the affinity was found to be 2 μM when binding to ME(pS)VD, compared to 67 μM for MESVD for the best TRAP variant, but the affinity to MEEVD was not reported. This interaction was also demonstrated intracellularly by fusing the peptide or the TRAP to either half of a split-mCherry detection system, which resulted in a reconstitution of fluorescence.

Based on the first redesigns (Jackrel et al., 2009) to recognize a different the C-terminal peptide residue, Speltz et al. (2015b) took a combined approach including rational design and directed evolution to create three new TRAPs, binding to either MEEVV, MERVW or MRRVW with micromolar affinity. This included the creation of both a peptide library and a pocket library in the protein, which were screened against each other using either a split-GFP assay or an *in-vitro* pull-down assay. TRAP1 (MEEVV)

also showed some cross-reactivity by binding to the peptide MERVW, but still less well than to its cognate target. Having these new target-TRAP pairs in hand, the authors showed that the affinity is gradually decreased when changing the C-terminal tryptophan to smaller hydrophobic residues (leucine and isoleucine).

Several applications of the CTPR scaffold have recently been published and reviewed, for example the use of TRAPs as an affinity column or as a biosensor by derivatizing gold nanoparticles with CTPRs while retaining the CTPR's ability to bind their target peptide (Couleaud et al., 2015; Speltz et al., 2015a). A general feature that is often observed in crystal structures is the head-to-tail and side-to-side interaction of repeat proteins with themselves. In the case of CTPRs this has been used to generate microfibrils, based on the head-to-tail stacking or achieved by linking via disulfides (see, e.g. Mejías et al., 2014).

CTPRs are an example of engineered repeat proteins that have been modified to bind to different amino acids of an unstructured, elongated peptide sequence in a perpendicular orientation with respect to the long axis of CTPR (Figure 2A). A different binding mode is found in ArmRPs, which bind their peptide in an antiparallel way and shall be discussed in the next section.

Armadillo repeat proteins

The family of ArmRPs share a common fold that is made up of repeating modules consisting of three helices, of which the longest, helix 3, is involved in binding to the target. As reviewed previously (Reichen et al., 2014a), designed ArmRPs (dArmRPs) based on the natural scaffold were successfully designed (Parmeggiani et al., 2008; Alfarano et al., 2012; Madhurantakam et al., 2012; Varadamsetty et al., 2012; Reichen et al., 2014a) and a first binder against the neuropeptide neurotensin was selected by ribosome display from a large 'universal' library (Varadamsetty et al., 2012). These engineered ArmRPs are built from 42 amino acid long repeats, which are derived from the natural fold of three consecutive helices per repeat. The internal repeats contributing to peptide binding form a hydrophobic core that is shielded by special capping repeats at both the N- and C-terminus. To analyze the complex structure and dynamics of this binder by nuclear magnetic resonance (NMR) a new strategy had to be developed (Ewald et al., 2015) since the assignment of individual amino acids was found difficult due to the repetitiveness, and in this case, the binding was additionally rather dynamic. First, a reduced-size binder was constructed that included only the repeats contributing to binding. With this binder and a multidisciplinary approach, which further included the use of molecular dynamics, docking procedures, chemical shift perturbations and paramagnetic relaxation enhancement, amino acid side chains could be assigned that were involved in binding the peptide. Having the protocol established, a spin label was used to analyze the structural contributions of the N-cap mutations that had been introduced previously (Alfarano et al., 2012) to stabilize the N-cap.

In another study, the dArmRPs were split in between two internal repeats (Watson et al., 2014), and a high tendency of the split proteins to self-complement was discovered, thereby forming a continuous fold with a shared hydrophobic core. NMR studies then showed that the halves come together in an orientation to almost perfectly regenerate the original protein. This study not only describes a potential evolutionary pathway of how repeat proteins may have arisen, but it also presents a useful strategy for the segmental labeling of dArmRPs, directly at the protein level.

Besides NMR, crystallography is of great use for supporting the design cycles, as it provides feedback with atomic resolution. Nonetheless, it was demonstrated in two studies (Reichen et al., 2014b, 2016) that the structures and the binding behavior of dArmRP can be influenced by the crystallization conditions. Calcium binding sites were found in the backbone of the scaffold, and binding of Ca^{2+} from the crystallization solution can influence the inter-repeat interactions of the protein and can even lead to a rigid-body movement of the C-cap, and thus influence the superhelical parameters. These studies show that for closing the design cycle by determining atomic resolution structures, the development of a reliable crystallization system that closely mimics the equilibrium solution structure in every detail is very important, and actually rather challenging, but such experiments are currently under way.

The ability of dArmRPs to bind a $(KR)_5$-peptide in a modular way with low nanomolar to even picomolar affinities was recently demonstrated (Hansen et al., 2016). By using a newly developed fluorescence anisotropy assay for the rapid determination of dissociation constants it was shown that individual arginines and lysines contribute regularly to binding, interacting with their respective pocket, with the binding energy provided by an arginine pocket being larger than that of a lysine pocket. Consequently, each of these identical protein repeats contributed the same energetic increment to binding. A crystal structure of a dArmRP with five internal repeats binding to a $(KR)_5$-peptide (Figure 2B) confirmed that the peptide

binds to the dArmRP in a modular way. Thus, in this system, each protein repeat contacts two side chains of the peptide with very regular energetic contributions, as well as the peptide main chain, thus laying the foundation for achieving such regular binding also for other peptide sequences by engineering and evolving the pockets in the protein appropriately.

Discussion and future aspects

Antibodies and other scaffold proteins have been selected to bind peptides with a variety of different sequences (Reichen et al., 2014a), with a wide range of affinities and specificities. Nevertheless, at the moment we are still limited to an individual selection for each new target peptide, in the same way as for every protein target. In other words, the modular nature of a peptide is not exploited at all.

Designed peptide-binding repeat proteins may provide a solution to this problem. Because of the perpendicular binding to the repeat protein axis, TPRs are so far limited to pentapeptides (see Figure 2A). A prerequisite for modular binding is that the peptide direction has to be antiparallel (or parallel) to the binding scaffold (Figure 2B), because only this allows an elongation of the binder with increasing length of the bound peptide. ArmRPs and their peptides have been engineered to different lengths, and even picomolar dissociation constants were described (Hansen et al., 2016). Furthermore, ArmRPs have shown binding to at least 10 amino acids in a modular way (Hansen et al., 2016).

Repeat proteins are highly suited for modular binding, as they already have a modular architecture. In the case of dArmRPs the peptide and the dArmRP fold into two superhelices, with the ArmRP winding around the peptide helix, with every peptide-binding module making contact to two amino acids (see Figure 2B).

The considerable protein engineering challenge is to keep this modularity for as many different peptide sequences as possible. Even small inaccuracies may add up with longer peptides and lead to a decrease in modular binding, a problem that might be addressed by slightly flexible scaffolds. At present it is not clear, however, how big the discrimination power between similar amino acids will be.

The described challenges of modular peptide binding – keeping geometric regularity and thus specificity and affinity over a sufficient length – could make it necessary to design new scaffolds that are not based on natural proteins. These new scaffolds could provide a wider range of different pockets for different amino acids or engineer the binding geometry over a wider range depending on the peptide sequence. Thereby, a more accurate design of the inter-pocket distances could be achieved, thereby allowing longer peptides of arbitrary sequence to be bound, by exploiting new curvatures and geometries, which are better suited to bind elongated peptides of particular sequence (see, e.g., Figure 2C). While peptide binding has not yet been described for such new scaffolds, the following section should give some examples of what has been achieved in the last few years with regards to the computational design of repeat proteins.

Classical repeat protein designs started from natural repeat proteins and often used a sequence-based consensus approach, sometimes combined with computational design. With the growing number of sequences and better computational design strategies, new and stable repeat protein scaffolds can now be generated, not only based on known structures but also completely *de novo* (reviewed, e.g. in Woolfson et al., 2015). Examples of leucine-rich repeats with defined curvatures have been derived using combinations of consensus design and computational methods (see, e.g. Park et al., 2015). Here we want to concentrate on recent *de novo* designs of α-solenoids, even though no binding has been reported yet.

One example to guide repeat protein design was presented recently (Sawyer et al., 2013). Repeat protein modules were not defined as single repeats but rather as conserved motifs which can include more than one repeat. The different annotation revealed new inter-repeat cross-correlations and patterns, which can help in detecting additional conserved positions that might have been overlooked by a simple one-repeat consensus design.

A new approach was recently reported that combines not only sequence data and structural information but also relies on the use of Rosetta *de novo* design, aiming to avoid model bias (Parmeggiani et al., 2015). This can be a problem in cases where there is a lack of a sufficient number of input structures. The approach has resulted in new backbone designs which included also ArmRPs and TPRs.

In contrast to sequence-based designs, fully reference-free methods might reveal tertiary structures that nature has not developed. The already mentioned Rosetta *de novo* design method was used to assess whether new α-solenoid folds can be designed from scratch (Doyle et al., 2015). By predefining only geometric parameters and concentrating on left-handed designs that are rarely found in nature, new repeat proteins could

be generated. These designs were highly stable and showed characteristics found in other repeat proteins, for example tubular assemblies and head-to-tail interactions in crystal structures.

Doyle et al. (2015) thus revealed that existing repeat proteins cover only a very small fraction of possible folds. Another very systematic way to design new folds took the simple helix-turn-helix motif as a starting point. By varying the lengths of the helices and linkers, a fully automatic and very precise algorithm was employed for the design (Brunette et al., 2015) (example structures are shown in Figure 2C). For 14 out of 15 solved crystal structures the initial designs could be used to perform molecular replacement directly. The two latter studies show the power of current computational methods. It will now be interesting to extend such approaches to the generation of binding proteins to comparing them as a starting points to consensus-derived scaffolds.

The discussed design strategies and scaffolds provide a variety of future alternatives not only for the design of modular peptide binders. The developed techniques could also serve for the improvement of existing scaffolds or lay the basis for new classes of binding proteins.

Conclusions

Recent studies on TPRs and dArmRPs have shown that these are well-suited scaffolds for the design of new peptide binders, where dArmRPs repeats are colinear with the dipeptide repeats in the bound peptide. The examples of designed repeat proteins mentioned demonstrate that in terms of scaffold design we are not limited to natural folds anymore. Future experiments will have to test whether artificial repeat proteins can be used for creating new binding proteins, since up to now all protein-peptide interactions have been largely inspired by naturally occurring binding pockets.

Funding: Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (Grant/Award Number: CRSII3_141832, Sinergia Program).

References

Alfarano, P., Varadamsetty, G., Ewald, C., Parmeggiani, F., Pellarin, R., Zerbe, O., Plückthun, A., and Caflisch, A. (2012). Optimization of designed armadillo repeat proteins by molecular dynamics simulations and NMR spectroscopy. Protein Sci. *21*, 1298–1314.

Bradbury, A. and Plückthun, A. (2015). Reproducibility: standardize antibodies used in research. Nature *518*, 27–29.

Brunette, T., Parmeggiani, F., Huang, P.S., Bhabha, G., Ekiert, D.C., Tsutakawa, S.E., Hura, G.L., Tainer, J.A., and Baker, D. (2015). Exploring the repeat protein universe through computational protein design. Nature *528*, 580–584.

Cohen, S.S., Riven, I., Cortajarena, A.L., Rosa, L. De, D'Andrea, L.D., Regan, L., and Haran, G. (2015). Probing the molecular origin of native-state flexibility in repeat proteins. J. Am. Chem. Soc. *137*, 10367–10373.

Cortajarena, A.L., Kajander, T., Pan, W., Cocco, M.J., and Regan, L. (2004). Protein design to understand peptide ligand recognition by tetratricopeptide repeat proteins. Protein Eng. Des. Sel. *17*, 399–409.

Cortajarena, A.L., Wang, J., and Regan, L. (2010). Crystal structure of a designed tetratricopeptide repeat module in complex with its peptide ligand. FEBS J. *277*, 1058–1066.

Couleaud, P., Adan-Bermudez, S., Aires, A., Mejías, S.H., Sot, B., Somoza, A., and Cortajarena, A.L. (2015). Designed modular proteins as scaffolds to stabilize fluorescent nanoclusters. Biomacromolecules *16*, 3836–3844.

Doyle, L., Hallinan, J., Bolduc, J., Parmeggiani, F., Baker, D., Stoddard, B.L., and Bradley, P. (2015). Rational design of α-helical tandem repeat proteins with closed architectures. Nature *528*, 585–588.

Ewald, C., Christen, M.T., Watson, R.P., Mihajlovic, M., Zhou, T., Honegger, A., Plückthun, A., Caflisch, A., and Zerbe, O. (2015). A combined NMR and computational approach to investigate peptide binding to a designed armadillo repeat protein. J. Mol. Biol. *427*, 1916–1933.

Hansen, S., Tremmel, D., Madhurantakam, C., Reichen, C., Mittl, P.R.E., and Plückthun, A. (2016). Structure and energetic contributions of a designed modular peptide-binding protein with picomolar affinity. J. Am. Chem. Soc. *138*, 3526–3532.

Jackrel, M.E., Valverde, R., and Regan, L. (2009). Redesign of a protein-peptide interaction; Characterization and applications. Protein Sci. *18*, 762–774.

Jost, C. and Plückthun, A. (2014). Engineered proteins with desired specificity: DARPins, other alternative scaffolds and bispecific IgGs. Curr. Opin. Struct. Biol. *27*, 102–112.

Madhurantakam, C., Varadamsetty, G., Grütter, M.G., Plückthun, A., and Mittl, P.R.E. (2012). Structure-based optimization of designed armadillo-repeat proteins. Protein Sci. *21*, 1015–1028.

Mejías, S.H., Sot, B., Guantes, R., and Cortajarena, A.L. (2014). Controlled nanometric fibers of self-assembled designed protein scaffolds. Nanoscale *6*, 10982–10988.

Millership, C., Phillips, J.J., and Main, E.R.G. (2016). Ising model reprogramming of a repeat protein's equilibrium unfolding pathway. J. Mol. Biol. *428*, 1804–1817.

Orlicky, S., Tang, X., Willems, A., Tyers, M., and Sicheri, F. (2003). Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase. Cell *112*, 243–256.

Park, K., Shen, B.W., Parmeggiani, F., Huang, P.S., Stoddard, B.L., and Baker, D. (2015). Control of repeat-protein curvature by computational protein design. Nat. Struct. Mol. Biol. *22*, 167–174.

Parmeggiani, F., Huang, P.S., Vorobiev, S., Xiao, R., Park, K., Caprari, S., Su, M., Seetharaman, J., Mao, L., Janjua, H., et al. (2015). A general computational approach for repeat protein design. J. Mol. Biol. *427*, 563–575.

Parmeggiani, F., Pellarin, R., Larsen, A.P., Varadamsetty, G., Stumpp, M.T., Zerbe, O., Caflisch, A., and Plückthun, A. (2008). Designed armadillo repeat proteins as general peptide-binding scaffolds: consensus design and computational optimization of the hydrophobic core. J. Mol. Biol. *376*, 1282–1304.

Petsalaki, E. and Russell, R.B. (2008). Peptide-mediated interactions in biological systems: new discoveries and applications. Curr. Opin. Biotechnol. *19*, 344–350.

Plückthun, A. (2015). Designed ankyrin repeat proteins (DARPins): binding proteins for research, diagnostics, and therapy. Annu. Rev. Pharmacol. Toxicol. *55*, 489–511.

Reichen, C., Hansen, S., and Plückthun, A. (2014a). Modular peptide binding: from a comparison of natural binders to designed armadillo repeat proteins. J. Struct. Biol. *185*, 147–162.

Reichen, C., Madhurantakam, C., Plückthun, A., and Mittl, P.R.E. (2014b). Crystal structures of designed armadillo repeat proteins: implications of construct design and crystallization conditions on overall structure. Protein Sci. *23*, 1572–1583.

Reichen, C., Madhurantakam, C., Hansen, S., Grütter, M.G., Plückthun, A., and Mittl, P.R.E. (2016). Structures of designed armadillo-repeat proteins show propagation of inter-repeat interface effects. Acta Crystallogr. Sect. D, Struct. Biol. *72*, 168–175.

Sawyer, N., Chen, J., and Regan, L. (2013). All repeats are not equal: a module-based approach to guide repeat protein design. J. Mol. Biol. *425*, 1826–1838.

Sawyer, N., Gassaway, B.M., Haimovich, A.D., Isaacs, F.J., Rinehart, J., and Regan, L. (2014). Designed phosphoprotein recognition in *Escherichia coli*. ACS Chem. Biol. *9*, 2502–2507.

Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., Hartl, F.U., and Moarefi, I. (2000). Structure of TPR domain-peptide complexes: Critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. Cell *101*, 199–210.

Speltz, E.B., Brown, R.S.H., Hajare, H.S., Schlieker, C., and Regan, L. (2015a). A designed repeat protein as an affinity capture reagent. Biochem. Soc. Trans. *43*, 874–880.

Speltz, E.B., Nathan, A., and Regan, L. (2015b). Design of protein-peptide interaction modules for assembling supramolecular structures *in vivo* and *in vitro*. ACS Chem. Biol. *10*, 2108–2115.

Urvoas, A., Guellouz, A., Valerio-Lepiniec, M., Graille, M., Durand, D., Desravines, D.C., Tilbeurgh, H. van, Desmadril, M., and Minard, P. (2010). Design, production and molecular structure of a new family of artificial α-helicoidal repeat proteins (αRep) based on thermostable HEAT-like repeats. J. Mol. Biol. *404*, 307–327.

Varadamsetty, G., Tremmel, D., Hansen, S., Parmeggiani, F., and Plückthun, A. (2012). Designed armadillo repeat proteins: library generation, characterization and selection of peptide binders with high specificity. J. Mol. Biol. *424*, 68–87.

Watson, R.P., Christen, M.T., Ewald, C., Bumbak, F., Reichen, C., Mihajlovic, M., Schmidt, E., Güntert, P., Caflisch, A., Plückthun, A., et al. (2014). Spontaneous self-assembly of engineered armadillo repeat protein fragments into a folded structure. Structure *22*, 985–995.

Woolfson, D.N., Bartlett, G.J., Burton, A.J., Heal, J.W., Niitsu, A., Thomson, A.R., and Wood, C.W. (2015). *De novo* protein design: how do we expand into the universe of possible protein structures? Curr. Opin. Struct. Biol. *33*, 16–26.