Repeat protein engineering: creating functional nanostructures/biomaterials from modular building blocks.

Main, ER; Phillips, JJ; Millership, C

Biochemical Society TRANSACTIONS
www.biochemsoctrans.org
Biochemical Society Transactions

Repeat protein engineering: creating functional nanostructures/biomaterials from modular building blocks

Ewan R.G. Main*[1], Jonathan J. Phillips*[2] and Charlotte Millership*
*School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, U.K.

Abstract

There is enormous interest in molecular self-assembly and the development of biological systems to form smart nanostructures for biotechnology (so-called 'bottom-up fabrications'). Repeat proteins are ideal choices for development of such systems as they: (i) possess a relatively simple relationship between sequence, structure and function; (ii) are modular and non-globular in structure; (iii) act as diverse scaffolds for the mediation of a diverse range of protein–protein interactions; and (iv) have been extensively studied and successfully engineered and designed. In the present review, we summarize recent advances in the use of engineered repeat proteins in the self-assembly of novel materials, nanostructures and biosensors. In particular, we show that repeat proteins are excellent monomeric programmable building blocks that can be triggered to associate into a range of morphologies and can readily be engineered as stimuli-responsive biofunctional materials.

Introduction

Repeat or solenoid proteins are incredibly ubiquitous protein folds that typically form non-globular three-dimensional structures. They are present in nearly all forms of life and are involved in a myriad of essential processes. This repertoire derives from a modular construction that is produced from the stacking of tandemly arranged small structural motifs (20–40 residues) (Figure 1). They are mostly found in arrays of at least three repeats and can form the whole protein or simply be one domain of many [1–3]. Crucially, the modular nature of repeat proteins means that their structures are composed exclusively of regularized interactions that are local in primary sequence, contrasting with the complex topologies of globular proteins that are stabilized by many interactions distant with regard to the primary sequence. Moreover, the modular architecture of repeat proteins seems to have given them an evolutionary advantage over globular proteins as they can evolve via both point mutations and insertion/deletion/reshuffling of whole repeat motifs. This capacity has enabled them to evolve into a diverse range of protein–protein interaction scaffolds and have functions as varied as biological diffraction gratings and mechanical nanosprings [3–8].

The combination of modularity, non-globular fold, simplistic architecture and protein–protein interaction scaffold has made repeat proteins extremely popular systems to engineer and design (some examples include [1–3,9–25]). Thus, in the last 10 years, there have been many advances with highlights that include: (i) designing hyper-stable repeat proteins [3,9,11,12,20,21,23], (ii) engineering these designed scaffolds to have novel protein or DNA-binding capabilities (antibody-like entities) [10,25–27], and (iii) showing that repeat proteins can be altered not only by single amino substitution, but also by the addition or deletion of whole repeats [2,19,28–30]. Interestingly, these features seem more unique to repeat proteins than to their globular counterparts and have been well documented with a number of excellent review articles (some examples include [1,2,10,31,32]). Excitingly, the success of these studies has initiated their development as smart nanostructures/novel biomaterials for biotechnology (so-called 'bottom-up fabrications'). Therefore, in the present review, we briefly recap relevant repeat protein engineering and design studies and then describe the emerging use of engineered repeat proteins in the self-assembly of novel materials, nanostructures and biosensors.

Repeat proteins: predictable, programmable and modular building blocks

To date, the most studied repeat proteins consist of five folds: the mostly α-helical ANK (ankyrin repeat), TPR (tetratricopeptide repeat), LRR (leucine-rich repeat), TAL (transcription activator-like) effectors and the all-β-helical β-roll motif (Figure 1). In each case, the folded repeat proteins form linear structures. Importantly, successfully designed proteins have been produced for each of these folds

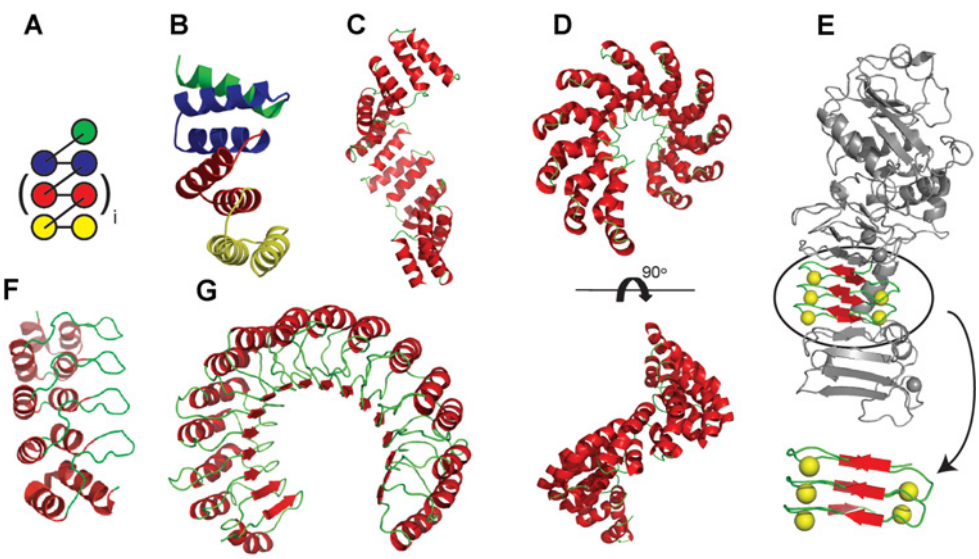
Key words: β-roll motif, engineering biomaterial, protein design, repeat protein, tetratricopeptide repeat.

Abbreviations used: Aha, azido-homoalanine; ANK, ankyrin repeat; CTPR, consensus tetratricopeptide repeat module; CTPRΔS, consensus tetratricopeptide repeat module with deleted C-terminal helix; LRR, leucine-rich repeat; NCL, native chemical ligation; TPR, tetratricopeptide repeat.

[1]**To whom correspondence should be addressed (email e.main@qmul.ac.uk).**
[2]**Present address:** MedImmune, Analytical Biochemistry and Antibody Discovery Protein Engineering, Granta Park, Cambridge CB21 6GH, U.K.

Figure 1 | **Topological and ribbon representations of repeat proteins**

(**A**) Topological map of a CTPR protein. The parentheses around the central TPR motif indicate that repeat proteins can be increased or decreased in size while retaining the same topology. For example, for CTPR3, $i = 1$. (**B** and **C**) The crystal structures of CTPR3 (PDB code 1NA0) and CTPRa8 (PDB code 2FO7) respectively. In (**B**) the consecutive repeat units are coloured from yellow (N-terminus) to blue (C-terminus), with the C-terminal solvating helix (S) in green (absent from the crystal structure of CTPRa8). (**D**) The crystal structure of the TAL effector DNA-binding domain of dHax3 in two orientations (PDB code 3V6P). (**E**) The crystal structure of the alkaline protease of *Pseudomonas aeruginosa*. This is a two-domain protein with a Ca^{2+}-binding parallel β-roll motif circled and shown in colour (PDB code 1KAP). (**F**) The crystal structure of a designed ANK repeat protein (PDB code 2XEE). (**G**) The crystal structure of human placental RNase inhibitor (hRI), a LRR protein (PDB code 1A4Y). All repeat proteins are coloured according to secondary structure (red for α-helices and β-strands, and green for loops).



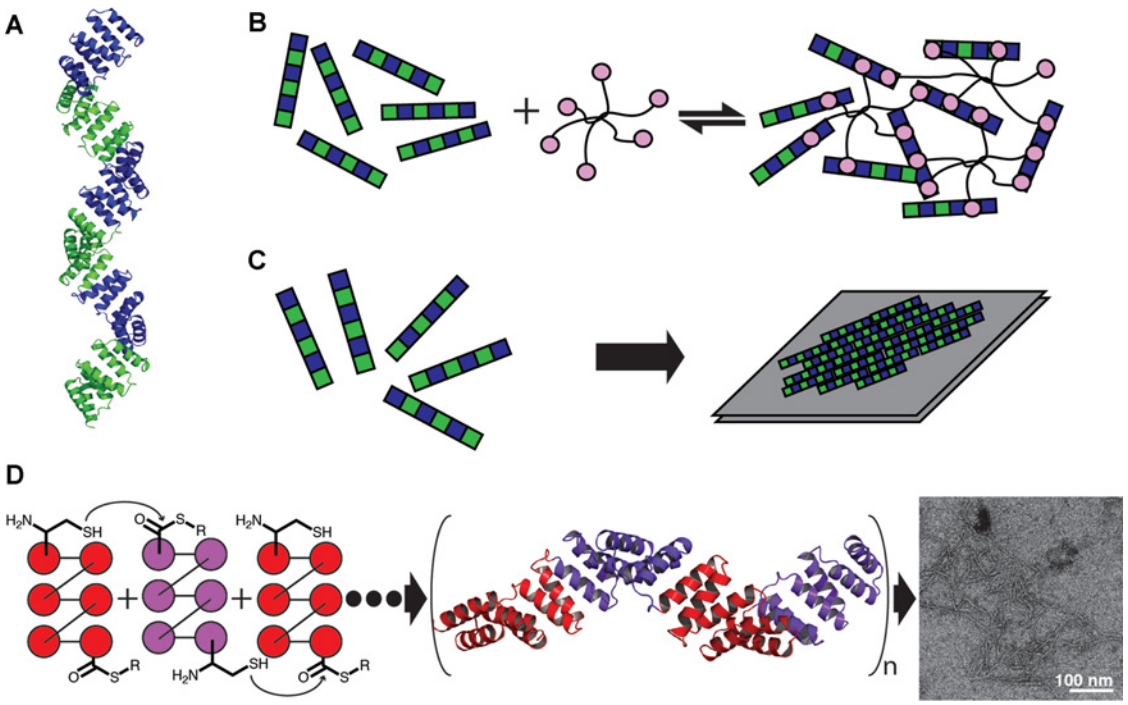
by constructing genes with various numbers of statistically driven consensus motifs or simple modification of an existing repeat motif [11,20,21,23,33]. These designs have shown that consensus proteins can be easily lengthened or shortened by the addition or removal of identical repeat motifs. Thus, as you increase the number of repeat motifs within a protein, large elongated superstructures with identical inter- and intra-repeat interactions are produced (Figure 1). Moreover, increasing the number of stacked repeats generates protein constructs with higher thermodynamic stability that can be modelled and predicted using a relatively simplistic Ising model approximation [13,19,28,34–37]. For example, proteins produced from CTPRs (consensus TPR modules) can be recombinantly produced to form highly thermodynamically stable very rigid rod-shaped superhelical arrays of up to 20 repeats [29,38]. The high thermostability of consensus scaffolds also enables them to be redesigned and engineered into novel binding proteins. Here, residues statistically shown to be hyper-variable and thus more likely to be important for binding specificity can be mutated either rationally or via directed evolution to produce proteins with specific high-affinity binding sites for a known target [10,24,32]. Thus repeat proteins should provide excellent building blocks for self-assembly as they have the ability to be combined in a modular fashion to create novel and scalable chimaeras/nanostructures with predictable structures, topologies, stabilities, linear length and function.

Creating self-assembled systems from repeated motifs

In order to create repeat protein biomaterials/biosensors, two main strategies have been adopted so far: (i) recombinant protein expression and purification of specific repeat designs followed by either spontaneous or triggered self-assembly into the desired biomaterial; and (ii) chemical synthesis of specific repeats coupled with an assembly trigger. Both strategies enable design of both the trigger to self-assembly and the building block modules used to form the resultant material. This has enabled proof-of-principle bare biomaterial scaffold production, followed by or with concurrent engineering in functionality and control of size and scale of the assemblies. Functionality is introduced through the incorporation of repeat modules that bind to specific ligands or introducing a stimulus or 'smart' response to an environmental cue by, for example, changing the physical state of the biomaterial through controlled triggering of either self-assembly or disassembly. To showcase how differing repeat scaffolds have been exploited to produce

Figure 2 | Triggered assembly of CTPRs to form hydrogels, films and fibrils

(**A**) Ribbon representation of CTPR18 protein, constructed from the crystal structure of CTPR8a (PDB code 2FO7). The CTPR18 is composed of alternating three-repeat peptide-binding modules and three-repeat spacer modules (coloured blue and green). (**B**) Schematic representation of CTPR18 (rectangles of alternating blue and green modules) combining with multivalent cognate peptide–PEG cross-linker (pink circles on black lines) to produce an interlinked hydrogel. (**C**) Schematic representation of CTPR18 (rectangles of alternating blue and green modules) forming a thick film after being deposited on a Teflon surface with a plasticizer and left to dry. (**D**) Scheme for NCL polymerization of CTPRΔS protein monomers. Orthogonal bifunctional protein monomers (shown as topological map) react in a head-to-tail fashion to form covalently bonded oligomers through NCL. These oligomers can continue to polymerize to form single protein chain fibrils of continuous CTPR superhelices (transmission EM micrograph negatively stained with uranyl acetate).



a range of differing types of biomaterials with various structural, functional and physical properties, we describe noteworthy examples of engineered repeat protein self-assembling biomaterials below.

TPR hydrogels, biofilms and fibrils

The α-helical TPR motif consists of a 34-amino-acid repeat that forms two antiparallel α-helices (Figures 1A–1C). These stack on top of each other and form superhelical structures with a complete superhelix requiring approximately eight repeats (Figures 1A–1C). Three studies have shown how recombinantly produced TPR building blocks can be used to create hydrogels [38], films [39] and fibrils [40]. In each case, differing triggers and/or building blocks were used. To create both the hydrogels and films, a functional and extremely rigid rod-like 18-mer CTPR protein was designed and produced. These were composed of alternating designed three-repeat non-functional 'spacer' modules and three-repeat peptide-binding modules [38] (Figure 2A). The CTPR18s were monomeric in solution, but could be triggered to create smart hydrogels or macroscopic solid films. To produce the hydrogels, the CTPR18 proteins were incubated with multivalent cognate peptide–PEG cross-linker. This caused multiple CTPR18s to be cross-linked together via the peptide–PEG component and form a gel (Figure 2B). The stimuli-responsive 'smart' element was encoded through the peptide-binding modules: upon introduction of salt, the gel disassembles due to disruption of the electrostatic peptide–protein interaction. Macroscopic solid films, on the other hand, were formed by subjecting a solution of CTPR18 with 1% (w/v) PEG 400 and leaving it to evaporate on the Teflon tape [39]. Interestingly, this produced a 100 μm multilayered film that retains its α-helical content and functional ability to bind cognate peptide (when generated in its presence). X-ray scattering, coupled with birefringence of the film, also suggested that the film possessed some macroscopic alignments and thus ordering of the CTPR18s (Figure 2C). The accompanying review by Regan et al. [41] in this issue of *Biochemical Society Transactions* discusses both of these studies in greater detail.

In comparison with the studies on hydrogel and film formation, we have explored techniques to produce repeat proteins of far greater size than recombinant expression and

purification can yield alone (such a process is limited, not least in the size and yield of protein that can be produced). Our work was inspired by the chemical syntheses of smaller sized repeat proteins and studies showing TPRs with compatible ends can be forced to self-associate through the addition of a cross-linker [42–44]. For example, chemical synthesis of CTPR3 was achieved by chemically producing individual repeats by solid-state peptide synthesis and assembly through a stepwise orthogonal protection with NCL (native chemical ligation) [44]. The prerequisites for NCL are a cysteine residue at the N-terminus and a thioesterified C-terminus. The NCL reaction proceeds through an attack of the thioester at the C-terminus by the N-terminal cysteine sulfur. An *N*-acyl rearrangement then takes place to yield a peptide bond linking both repeats together. Thus the ligated product is a polypeptide indistinguishable from one formed by other protein synthesis routes. To block intramolecular reactions, reacting peptides were synthesized with protecting groups covering orthogonal ends. To create larger assembles than either chemical synthesis or recombinant expression and purification alone, we genetically constructed a protein chimaera that contained a three-repeat CTPR3ΔS (consensus tetratricopeptide repeat module with deleted C-terminal helix) protein building block, with oligomerization-compatible N- and C-termini, flanked by an N-terminal polyhistidine tag and C-terminal intein domain. After recombinant production/purification, the C-terminal intein can be induced to self-cleave, leaving a thioester at the C-terminus by simply adding the reducing agent sodium 2-mercaptoethane sulfonate (MESNA). Interestingly and importantly for control of formation, the CTPR3ΔS building blocks remain monomeric even though their N- and C-termini have compatible interfaces. Moreover, their rigidly folded structure also prevents the building blocks from futile intraprotein cyclization. To trigger polymerization, Factor Xa protease is added (in mild conditions). This specifically cleaves the polyhistidine tag to reveal the N-terminal cysteine residue and thus yield an active monomer. NCL then occurred between the cysteine residue of one CTPR3ΔS with the C-terminal thioester of another (Figure 2D). Excitingly, this reaction proceeds smoothly with little non-specific aggregation to produce many specific oligomeric species within 48 h. Moreover, after a further 24 h, negatively stained transmission electron microscope images showed TPR fibrils. These polymers are consistent with single protein chains that are helically folded and show a filamentous morphology; the filaments exhibit high aspect ratio and extend up to microns in length. The individual filaments were also seen to associate into fibres and thereby may give rise to heterogeneous branching and, ultimately, extensive networks.

These three studies highlight the plasticity and range of morphologies that a single repeat fold can achieve. Using three distinct approaches the TPR scaffold was directed to form various levels of complexity from one-dimensional filaments to two-dimensional films and three-dimensional hydrogels. Each morphology lends itself to a particular set of applications; however, most significantly, this diversity of fabrication derives from relatively simple modification of soluble TPR monomers.

β-Roll hydrogels and fibrils

In contrast with the α-helical TPR motif, the β-roll motif consists of two short β-strands that are linked by two loops that turn approximately 90° (Figure 1E). In many instances they bind divalent metals (commonly Ca^{2+}) within the turn regions and, when tandemly arrayed, form a β-helical structure. Interestingly, if the divalent cations are removed they unfold and form disordered chains [45]. This response to divalent metal ion concentration has been harnessed in two studies to produce novel triggerable fibre formation and a smart hydrogel [46,47] (Figure 3).

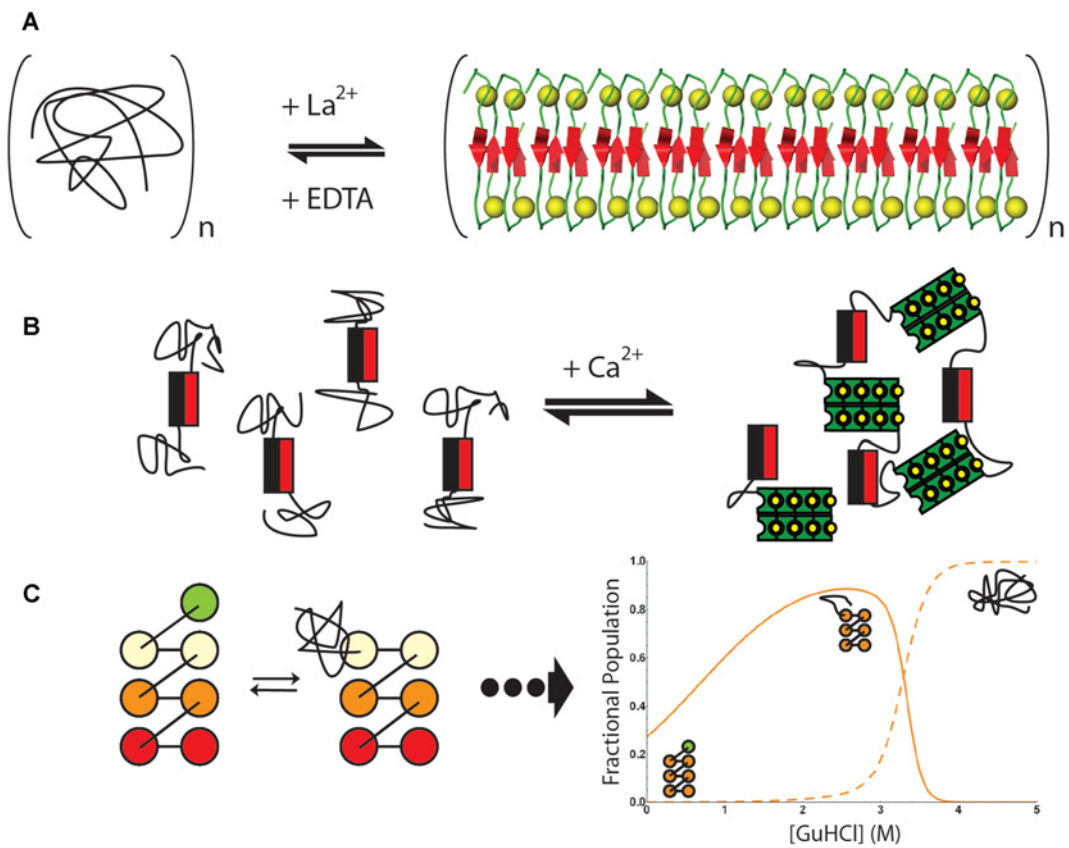
For fibre formation, Davies and co-workers designed minimized β-roll motifs that contained either three β-strands and two Ca^{2+}-binding sites (34 amino acids) or five β-strands and four Ca^{2+}-binding sites (50 amino acids) [47]. Originally these constructs had been designed to fold into monomeric scaffolds on addition of Ca^{2+}. However, after being produced using peptide synthesis, no structural change was observed on Ca^{2+} addition. However, when La^{2+} was added, filaments of approximately 3 nm in diameter and many nanometres in length were formed. In comparison with the covalent attachment of TPR molecules described in the previous section, the oligomerization of the β-roll arrays was driven through non-covalent metal-induced interactions (Figure 3A). Consequently, the metal-ion-mediated oligomerization was found to be reversible, as removing the La^{2+} through chelation with EDTA could disassemble the filaments. Some control of filament length was also obtained by introducing a further β-roll peptide that was capped at one end with biotin. In order to create a Ca^{2+}-responsive hydrogel, Banta and co-workers also used a β-roll-based protein [46], although here they created a chimaera that possessed a leucine zipper domain N-terminally attached to a modified β-roll array. The modified β-roll array was engineered to contain a leucine-rich face. Once recombinantly produced, the leucine zipper–β-roll chimaera form is soluble. However, when Ca^{2+} was added the β-roll array fold, exposing the leucine-rich face. This causes the β-roll arrays to interact and oligomerize. The β-roll array oligomerization, coupled with existing leucine zipper interactions, enabled a weak hydrogel to form (Figure 3B).

Future directions: new stimuli-responsive self-assembly?

One of the unique aspects that the above examples of repeat protein biomaterials demonstrate is that, unlike other systems (for example, amyloid and coiled coils), assembly must be triggered to occur rather than being an inherent property. Therefore one route to increasing the range of useful biosensing and biomaterial applications would be to diversify the repertoire of signals that can be transduced by repeat proteins

Figure 3 | Triggered assembly of β-rolls and a designed specific unfolding of a CTPR protein

(**A**) Scheme for reversible polymerization of the designed minimized β-roll motifs on addition of La^{2+}. When metal is absent, the designed minimized β-roll motifs are unfolded. Addition of La^{2+} (yellow balls) causes the β-rolls to become structured and polymerize. (**B**) Scheme for reversible gelation of a chimaera of a leucine zipper (black and red rectangles) with designed minimized β-roll motifs. On addition of Ca^{2+} (yellow circles) the minimized β-roll motifs fold (green squares) from unstructured polypeptides and oligomerize. (**C**) Schematic diagram of a designed variant of CTPR3 specifically unfolding to reveal an intermediate with an unfolded C-terminal α-helix (green circle) using the denaturant guanidinium chloride. The graph shows the simulated chemical denaturation for the mutant CTPR3 (broken line). The variant was designed by incorporating destabilizing mutations in the C-terminal capping helix. These reduce the helix's thermodynamic stability and the strength of its interaction with the rest of the protein. The result is a thermodynamic switch that causes an intermediate state with unfolded C-terminal helix to become highly populated at low guanidinium chloride concentrations (continuous line).



to trigger self-assembly. Interestingly, a number of studies show linking methods that could be adapted for this purpose [48]. Perhaps one of the most obvious is the use of Click chemistry, as it is orthogonal and results in high-yielding conjugation under mild conditions. Recently, Plückthun and co-workers have shown that designed ANK folds can be functionalized by introduction of clickable methionine surrogates such as Aha (azido-homoalanine) at the N-termini. After recombinant expression and purification of the construct, a PEG moiety was 'clicked' to the Aha N-terminus using the copper-free SPAAC (strain-promoted alkyne-azide cycloaddition) reaction [48]. Furthermore, Bertozzi and co-workers have demonstrated that copper-free Click chemistry is effective *in vivo*. They showed that cyclo-octyne reagents could be conjugated to cell-surface sialic acids in live mice [49]. This opens up the possibility of exploiting conjugation and signal transduction in physiologically relevant contexts.

More fundamentally, the inherent tuneability of repeat protein folding may also provide a path for further exploitation. It has been elegantly shown that both the kinetic and equilibrium repeat protein folding is prone to the population of partially folded intermediate states [1,2,13,29]. Moreover, by engineering changes in stability to individual repeats and/or the whole protein, both the kinetic folding pathway and population of the intermediate states can be changed [2,17,28,50]. Importantly, these studies open up the potential of exploiting the multistate folding of repeat proteins in the form of a switch, i.e. the engineered unfolding of part of a repeat protein (for example a single terminal repeat) to leave a folded intermediate unit with a

compatible oligomerization interface. We have explored this concept by using the predictive qualities of an Ising model to produce CTPR protein variants that thermodynamically unfold to a specific highly populated stable intermediate (C. Millership, J.J. Phillips and E.R.G. Main, unpublished work). This switch-like behaviour was designed by introducing asymmetry, via destabilizing mutations, into the C-terminus of a CTPR protein. The change in sequence causes the preferential unfolding of the terminal helix, uncoupling it from the rest of the protein, and reveals a stable intermediate with a potential dimerization interface (Figure 3C). The generality and extension of such an approach leads to exciting possibilities, for example temperature-sensitive protein switches. Furthermore, one can envisage designed repeat proteins as sensors that could oligomerize through a specifically triggered local unfolding event in response to external stimuli.

Conclusions

In the present review, we have briefly described the exciting new studies on the self-assembly of repeat proteins into stimuli-responsive biomaterials. Repeat proteins constitute ideal components for these applications as their modular structures enables production of monomeric programmable building blocks that can be triggered to associate into many morphological states. These properties will and have enabled the fabrication of triggerable novel biomaterials with desired bulk properties and with precise display of functional sites within the material.

Acknowledgements

We thank Dr R. Rose, Dr A. Lowe and members of School of Biological and Chemical Sciences (Queen Mary University of London) for a critical reading of the review and insightful discussions before submission.

Funding

C.M. is supported by a Biotechnology and Biological Sciences Research Council studentship.

References

1 Itzhaki, L.S. and Lowe, A.R. (2012) From artificial antibodies to nanosprings: the biophysical properties of repeat proteins. Adv. Exp. Med. Biol. **747**, 153–166
2 Main, E.R., Lowe, A.R., Mochrie, S.G., Jackson, S.E. and Regan, L. (2005) A recurring theme in protein engineering: the design, stability and folding of repeat proteins. Curr. Opin. Struct. Biol. **15**, 464–471
3 Main, E.R., Jackson, S.E. and Regan, L. (2003) The folding and design of repeat proteins: reaching a consensus. Curr. Opin. Struct. Biol. **13**, 482–489
4 Lee, G., Abdi, K., Jiang, Y., Michaely, P., Bennett, V. and Marszalek, P.E. (2006) Nanospring behaviour of ankyrin repeats. Nature **440**, 246–249
5 Kobe, B. and Kajava, A.V. (2001) The leucine-rich repeat as a protein recognition motif. Curr. Opin. Struct. Biol. **11**, 725–732
6 Groves, M.R. and Barford, D. (1999) Topological characteristics of helical repeat proteins. Curr. Opin. Struct. Biol. **9**, 383–389
7 Crookes, W.J., Ding, L.L., Huang, Q.L., Kimbell, J.R., Horwitz, J. and McFall-Ngai, M.J. (2004) Reflectins: the unusual proteins of squid reflective tissues. Science **303**, 235–238
8 Sedgwick, S.G. and Smerdon, S.J. (1999) The ankyrin repeat: a diversity of interactions on a common structural framework. Trends Biochem. Sci. **24**, 311–316
9 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
10 Tamaskovic, R., Simon, M., Stefan, N., Schwill, M. and Plückthun, A. (2012) Designed ankyrin repeat proteins (DARPins) from research to therapy. Methods Enzymol. **503**, 101–134
11 Lee, S.C., Park, K., Han, J., Lee, J.J., Kim, H.J., Hong, S., Heu, W., Kim, Y.J., Ha, J.S., Lee, S.G. et al. (2012) Design of a binding scaffold based on variable lymphocyte receptors of jawless vertebrates by module engineering. Proc. Natl. Acad. Sci. U.S.A. **109**, 3299–3304
12 Urvoas, A., Guellouz, A., Valerio-Lepiniec, M., Graille, M., Durand, D., Desravines, D.C., van Tilbeurgh, H., 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
13 Javadi, Y. and Main, E.R. (2009) Exploring the folding energy landscape of a series of designed consensus tetratricopeptide repeat proteins. Proc. Natl. Acad. Sci. U.S.A. **106**, 17383–17388
14 Tripp, K.W. and Barrick, D. (2008) Rerouting the folding pathway of the Notch ankyrin domain by reshaping the energy landscape. J. Am. Chem. Soc. **130**, 5681–5688
15 Steiner, D., Forrer, P. and Plückthun, A. (2008) Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display. J. Mol. Biol. **382**, 1211–1227
16 Courtemanche, N. and Barrick, D. (2008) The leucine-rich repeat domain of Internalin B folds along a polarized N-terminal pathway. Structure **16**, 705–714
17 Lowe, A.R. and Itzhaki, L.S. (2007) Rational redesign of the folding pathway of a modular protein. Proc. Natl. Acad. Sci. U.S.A. **104**, 2679–2684
18 Main, E.R., Stott, K., Jackson, S.E. and Regan, L. (2005) Local and long-range stability in tandemly arrayed tetratricopeptide repeats. Proc. Natl. Acad. Sci. U.S.A. **102**, 5721–5726
19 Kajander, T., Cortajarena, A.L., Main, E.R., Mochrie, S.G. and Regan, L. (2005) A new folding paradigm for repeat proteins. J. Am. Chem. Soc. **127**, 10188–10190
20 Main, E.R., Xiong, Y., Cocco, M.J., D'Andrea, L. and Regan, L. (2003) Design of stable α-helical arrays from an idealized TPR motif. Structure **11**, 497–508
21 Kohl, A., Binz, H.K., Forrer, P., Stumpp, M.T., Plückthun, A. and Grutter, M.G. (2003) Designed to be stable: crystal structure of a consensus ankyrin repeat protein. Proc. Natl. Acad. Sci. U.S.A. **100**, 1700–1705
22 Forrer, P., Stumpp, M.T., Binz, H.K. and Plückthun, A. (2003) A novel strategy to design binding molecules harnessing the modular nature of repeat proteins. FEBS Lett. **539**, 2–6
23 Mosavi, L.K., Minor, D.L. and Peng, Z.Y. (2002) Consensus-derived structural determinants of the ankyrin repeat motif. Proc. Natl. Acad. Sci. U.S.A. **99**, 16029–16034
24 Cortajarena, A.L., Liu, T.Y., Hochstrasser, M. and Regan, L. (2010) Designed proteins to modulate cellular networks. ACS Chem. Biol. **5**, 545–552
25 Cortajarena, A.L., Yi, F. and Regan, L. (2008) Designed TPR modules as novel anticancer agents. ACS Chem. Biol. **3**, 161–166
26 Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J. and Voytas, D.F. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. **39**, e82
27 Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science **326**, 1509–1512
28 Phillips, J.J., Javadi, Y., Millership, C. and Main, E.R. (2012) Modulation of the multistate folding of designed TPR proteins through intrinsic and extrinsic factors. Protein Sci. **21**, 327–338
29 Cortajarena, A.L. and Regan, L. (2011) Calorimetric study of a series of designed repeat proteins: modular structure and modular folding. Protein Sci. **20**, 336–340

30 Tripp, K.W. and Barrick, D. (2004) The tolerance of a modular protein to duplication and deletion of internal repeats. J. Mol. Biol. **344**, 169–178
31 Kloss, E., Courtemanche, N. and Barrick, D. (2008) Repeat-protein folding: new insights into origins of cooperativity, stability, and topology. Arch. Biochem. Biophys. **469**, 83–99
32 Grove, T.Z., Cortajarena, A.L. and Regan, L. (2008) Ligand binding by repeat proteins: natural and designed. Curr. Opin. Struct. Biol. **18**, 507–515
33 Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G.M. and Arlotta, P. (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat. Biotechnol. **29**, 149–153
34 Aksel, T., Majumdar, A. and Barrick, D. (2011) The contribution of entropy, enthalpy, and hydrophobic desolvation to cooperativity in repeat-protein folding. Structure **19**, 349–360
35 Aksel, T. and Barrick, D. (2009) Analysis of repeat-protein folding using nearest-neighbor statistical mechanical models. Methods Enzymol. **455**, 95–125
36 Cortajarena, A.L., Mochrie, S.G. and Regan, L. (2008) Mapping the energy landscape of repeat proteins using NMR-detected hydrogen exchange. J. Mol. Biol. **379**, 617–626
37 Mello, C.C. and Barrick, D. (2004) An experimentally determined protein folding energy landscape. Proc. Natl. Acad. Sci. U.S.A. **101**, 14102–14107
38 Grove, T.Z., Osuji, C.O., Forster, J.D., Dufresne, E.R. and Regan, L. (2010) Stimuli-responsive smart gels realized via modular protein design. J. Am. Chem. Soc. **132**, 14024–14026
39 Grove, T.Z., Regan, L. and Cortajarena, A.L. (2013) Nanostructured functional films from engineered repeat proteins. J. R. Soc. Interface **10**, 20130051
40 Phillips, J.J., Millership, C. and Main, E.R. (2012) Fibrous nanostructures from the self-assembly of designed repeat protein modules. Angew. Chem. Int. Ed. **51**, 13132–13135
41 Sawyer, N., Speltz, E.B. and Regan, L. (2013) NextGen protein design. Biochem. Soc. Trans. **41**, 1131–1136
42 Eisenberg, M., Shumacher, I., Cohen-Luria, R. and Ashkenasy, G. (2013) Dynamic combinatorial libraries of artificial repeat proteins. Bioorg. Med. Chem. **21**, 3450–3457
43 Krachler, A.M., Sharma, A. and Kleanthous, C. (2010) Self-association of TPR domains: lessons learned from a designed, consensus-based TPR oligomer. Proteins **78**, 2131–2143
44 Bang, D. and Kent, S.B. (2005) His_6 tag-assisted chemical protein synthesis. Proc. Natl. Acad. Sci. U.S.A. **102**, 5014–5019
45 Chenal, A., Guijarro, J.I., Raynal, B., Delepierre, M. and Ladant, D. (2009) RTX calcium binding motifs are intrinsically disordered in the absence of calcium: implication for protein secretion. J. Biol. Chem. **284**, 1781–1789
46 Dooley, K., Kim, Y.H., Lu, H.D., Tu, R. and Banta, S. (2012) Engineering of an environmentally responsive β roll peptide for use as a calcium-dependent cross-linking domain for peptide hydrogel formation. Biomacromolecules **13**, 1758–1764
47 Scotter, A.J., Guo, M., Tomczak, M.M., Daley, M.E., Campbell, R.L., Oko, R.J., Bateman, D.A., Chakrabartty, A., Sykes, B.D. and Davies, P.L. (2007) Metal ion-dependent, reversible, protein filament formation by designed β-roll polypeptides. BMC Struct. Biol. **7**, 63
48 Simon, M., Zangemeister-Wittke, U. and Plückthun, A. (2012) Facile double-functionalization of designed ankyrin repeat proteins using Click and thiol chemistries. Bioconjugate Chem. **23**, 279–286
49 Chang, P.V., Prescher, J.A., Sletten, E.M., Baskin, J.M., Miller, I.A., Agard, N.J., Lo, A. and Bertozzi, C.R. (2010) Copper-free Click chemistry in living animals. Proc. Natl. Acad. Sci. U.S.A. **107**, 1821–1826
50 Werbeck, N.D. and Itzhaki, L.S. (2007) Probing a moving target with a plastic unfolding intermediate of an ankyrin-repeat protein. Proc. Natl. Acad. Sci. U.S.A. **104**, 7863–7868

Received 11 June 2013
doi:10.1042/BST20130102