ELSEVIER

Engineering and design
Protein design: theory and practice
Editorial overview

Lynne Regan and Sophie E Jackson

Current Opinion in Structural Biology 2003, **13**:479–481

0959-440X/$ – see front matter

DOI 10.1016/S0959-440X(03)00111-8

Lynne Regan

Department of Chemistry and Department of Molecular Biophysics & Biochemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520, USA
e-mail: lynne.regan@yale.edu

Sophie E Jackson

Cambridge University Chemical Laboratory, Cambridge University, Lensfield Road, Cambridge CB2 1EW, UK
e-mail: sej13@cam.ac.uk

We chose this set of reviews to give a broad coverage of the current state of protein engineering and design. The topics chosen encompass basic design approaches to study protein structure and function, comparisons of the strengths and limitations of screens and selections, and the application of engineering and design to proteins of therapeutic importance. It is a testament to the success of this field, and its future promise, that several engineered proteins are either approved or in trials for clinical applications.

In the area of basic protein design, Main, Jackson and Regan discuss recent successes in understanding the structure and folding of repeat proteins. These proteins have characteristics both in common with and disparate from globular proteins. Several successful idealised designs for repeat proteins have been reported, which provide a better understanding of their folding and stability. Repeat proteins are now being used as the framework for functional designs.

Metalloproteins were one of the first class of targets for functional protein design. The design and modification of metalloproteins is still very attractive because of the useful electronic, optoelectronic and catalytic properties that are associated with metal ion sites in proteins. Barker discusses two complementary approaches to the modification and design of metalloproteins: introducing novel metal-ion-binding sites into virgin proteins or manipulating the properties of natural metalloproteins — ''bottom-up'' versus ''top-down'' design. Successes include the design of novel haem-containing proteins and defining the range of redox potentials available to a natural metalloprotein. Barker concludes that ''evolution is a very creative inorganic chemist'' and that there are many ways to solve the problem, the best strategy to adopt being very dependent on the metal in the protein.

Genetic screens or selections for a phenotype of interest can be very powerful. Dalby discusses the use of directed evolution as a means to optimise enzymatic function. He discusses different possible strategies and their associated limitations. For example, one can chose either to mutate a small number of residues completely or to mutate the whole of a protein less completely. Why not mutate the whole of the protein completely? 20^{100} combinations for a 100-residue protein is clearly an unfeasibly huge number of variants! Dalby argues, however, that in targeting a few residues close to the 'active site' for more complete mutagenesis, one may completely miss more distal residues that contribute significantly.

Dalby describes some of the traditional approaches that have been employed to manipulate the catalytic fitness landscape and also new approaches that are overcoming some of the limitations of the traditional

methods. New work on 'SMART' libraries reduces the redundancy in normal libraries by using computational methods to preselect residues tolerant to mutagenesis. For β-lactamase, for example, a library of 7×10^{23} possible sequences was reduced to 172,800 and, from this SMART library, variants were selected with a 1300-fold increase in cefotaxime resistance. New genetic tools, including random deletion/insertion, exon shuffling and synthetic shuffling methods, represent new ways in which different types of library can be constructed. Increased screening and selection, the use of non-natural amino acids and smart libraries will all improve the ability to search through sequence/structure space. The rules learnt from such studies should lead to a better understanding of the relationship between sequence/structure and function, which will ultimately allow more rational design strategies and aid in optimising mutagenic strategies.

Typically, to investigate the role of a given amino acid, one mutates it to another naturally occurring amino acid. Such mutagenesis is a fairly blunt tool by which to dissect protein structure and function. The ability to introduce into proteins non-natural amino acids that have more finely tuned differences from the original sidechain would be extremely useful. Recently, the limited diversity of natural polypeptides has begun to be overcome by the incorporation of non-natural amino acids into peptide and protein combinatorial libraries. Roberts and co-workers outline the different strategies for expanding current combinatorial peptide and protein libraries, thereby increasing chemical complexity. The review focuses on libraries created using *in vitro* mRNA display technologies. Such technologies allow the use of nonsense suppression, chemical derivatisation, missense suppression and codon reassignment, in addition to very novel and ambitious projects to expand the genetic code using four-base codon–anticodon pairings.

The field has developed as a result of the union of ribosomal chemistry with modern synthetic methods and has been facilitated by the fact that the ribosome is a remarkably tolerant and versatile synthetic machine. It is clear that these techniques, which not only add diversity to libraries but also introduce many other desirable properties such as rigidity, will have an enormous impact on therapeutic and diagnostic ligand discovery.

Mayo, Desjarlais and colleagues discuss recent progress in the engineering and design of therapeutic proteins and peptides. In addition to the simple function of the protein or peptide to bind specifically to its target, all the other factors and properties that are prerequisites of a feasible therapeutic molecule, such as stability, bioavailability, solubility and low immunogenicity, as well as pharmacokinetics, pharmacodynamics and so on, are now considered early in any design or engineering project. Several success stories in the redesign of antibodies, cytokines and enzymes for therapeutic applications are described. The classical principles of protein design and modelling remain very important in the design of protein/peptide therapeutics; this is well illustrated by the design of a peptide inhibitor of membrane fusion targeted at the HIV viral fusion process involving gp41 and gp120. This is an excellent example by which knowledge gained from academic work on the design of four-helix bundle proteins proved invaluable in the design of the potentially therapeutic peptide.

Remaining on the theme of protein and peptide therapeutics, Presta reviews the current state of therapeutic antibodies. Although much of the original work in this field concentrated on humanising antibodies and novel antibody formats such as diabodies, there has been a recent expansion of engineering projects aimed at improving the efficacy of antibodies by a variety of methods. Many of these involve the formation of an antibody fusion of some type, for example, bispecific antibodies that bind two epitopes and can be constructed from full-length IgG or smaller fragments. In these cases, one arm targets specific cell surface receptors. They are designed to recruit immune system effector cells to kill tumour cells and have been shown to be highly effective. The engineering of antibody–cytokine fusions has also resulted in some recent success stories. Alternative strategies in which the antibody is linked, either chemically or genetically, to a toxin (originally to bacterial or plant toxins, but more recently to human toxins in order to obviate toxicity and immunogenicity problems with the former) are described.

One of the ultimate goals in protein design is, of course, to interfere with biological processes in order to generate new therapeutic agents. One biological process that is fast becoming the target of many design projects is protein deposition, such as the formation of amyloid fibrils. Now linked to many disease states, in particular those associated with old age, these amyloidoses involve the formation of toxic soluble oligomers and/or insoluble fibrils by the aggregation of otherwise soluble proteins. Doig and co-workers review some recent work in this highly active field. Strategies aimed at intervening at different steps in the amyloidosis process are being investigated, including inhibition of the expression of the relevant protein, stabilisation of the normal native state of the protein (which has been shown to work well for transthyretin), inhibition of the release of peptide if proteolysis is a key step (as is the case with Alzheimer's disease) or strategies aimed at inhibiting the aggregation process itself. Screening of small organic ligand libraries has met with some success, as has derivatisation of the wild-type peptide, such that binding but no extension of the amyloid fibril can occur. For Alzheimer's disease, this approach has been shown to be successful, at least at suppressing aggregation *in vitro* and reducing toxicity in

some cell lines. In particular, the core region of the Aβ peptide has been taken and engineered so that further oligomerisation can't take place. Various strategies have been employed, including the incorporation of proline residues, which cannot hydrogen bond through their backbone amide groups, and, in a similar manner, the use of *N*-methyl amino acids. It has recently become clear that it may not be the fibril itself that is toxic but a smaller soluble oligomer of the protein. This paves the way for alternative therapeutic strategies. A nice illustration of this is that a highly charged Aβ peptide actually accelerates aggregation but in doing so reduces toxicity, presumably by removing the toxic soluble oligomer rapidly from solution.

Together, these reviews provide an overview of the range and diversity of protein design and engineering successes during the past two years. We anticipate with optimism the exciting advances of the next two years, building upon these studies.