Engineering proteins: keeping up with biology

Donald F Doyle and David R Corey

A wide range of biological laboratories have adopted protein engineering techniques, altering the way biochemical research is carried out. Ironically, this broad success has increased the challenges faced by researchers at the chemistry–biology interface.

Address: Howard Hughes Medical Institute, Department of Pharmacology, University of Texas Southwestern Medical Center at Dallas, Dallas, TX 75235-9050, USA.

Correspondence: David R Corey
E-mail: corey@howie.swmed.edu

Chemistry & Biology July 1998, **5**:R157–R160
http://biomednet.com/elecref/10745521005R0157

 ISSN 1074-5521

The biologists' challenge to engineers

Few areas of biochemistry have intrigued as wide an audience as protein engineering. Manipulating proteins is intellectually fascinating and the potential to tailor function is a powerful attraction promising new experimental tools, industrial reagents, and novel therapeutics. Many of the successes of protein engineering over the last fifteen years have been achieved by researchers at the chemistry–biology interface motivated by a desire to achieve novel function by going beyond the constraints of naturally occurring polypeptide structures. Now, however, site-directed mutagenesis, phage display, chemical modification, and other tools for protein engineering have become widely available, which has led to the greatest success of protein engineering — its routine use by the general biology community.

For researchers who are specialists in particular areas of biological research, mutagenesis and protein-modification experiments are not protein engineering but simply tools that complement their existing expertise. Ironically, this almost unconscious use of protein engineering has produced results at least as remarkable as those produced by protein engineers themselves. One example is mutagenesis of adenylyl cyclase, a membrane-bound protein that contains two cytoplasmic domains and two membrane domains. As a membrane-bound protein the adenylyl cyclase was poorly expressed in recombinant form and difficult to characterize structurally. Tang and Gilman [1,2] reasoned that it might be possible to remove the membrane domains and to then couple the cytoplasmic domains using a linker region. There were many reasons to believe that this approach might not succeed, such as incorrect prediction of the boundaries of the membrane domains, misfolding of the variant protein or improper orientation of the cytoplasmic domains, but when the drastic remodeling was accomplished they found themselves in possession of a soluble adenylyl cyclase that retained most of the catalytic and regulatory properties of the full-length enzyme. This discovery has facilitated extensive kinetic analysis and has provided protein in a form that has allowed structure determination using X-ray crystallography [3].

The work by Tang and Gilman [1,2] and similar breakthroughs by other investigators sets a high standard because it combines general lessons for protein engineering with important basic insights within an established field. Such successes by laboratories that possess expert understanding of specific native macromolecules will become increasingly commonplace and present researchers at the chemistry/biology interface, who tend to be generalists, with a clear challenge — now that established biology laboratories with tremendous reservoirs of technical and scientific expertise in particular areas are comfortable with sophisticated engineering technologies how can researchers who lack these advantages compete? The danger is that a relative lack of expertise regarding a protein will lead the protein engineer to ask questions whose answers are not interesting to the biological audience ostensibly being addressed, leading biologists who specialize in the field to ask "What is really being learned?". That this question is asked reflects the different cultures of biology and chemistry — biologists are concerned with elucidating natural processes whereas chemists are accustomed to manipulating them. Given that a primary goal of chemical biology is to perform research that breaks new ground in biology, however, the question noted above must be considered and thoughtfully answered.

Why engineer proteins?

There are three generally applicable reasons to justify the resources devoted to engineering proteins. First, the engineered molecules could have practical uses as experimental tools or for the development and production of therapeutics. Second, during the process of engineering new techniques could be developed that will expand the options available for the research community as a whole. Third, by approaching a macromolecule from the perspective of an engineer, a new and unique understanding might be gained into how native macromolecules function. These reasons should be familiar as they are analogous to reasons for pursuing the synthesis of small molecules. Unlike synthetic organic chemistry, however, macromolecular engineering must achieve results that not only intrigue chemists but also provide important information or useful tools to biologists, thus answering the "what have you learned?" question noted above. We now

describe several recent examples of how researchers at the chemistry–biology interface have met this challenge.

Advantages of a protein engineer's perspective

Researchers with degrees in chemically related fields are accustomed to the concept of building and modifying molecules through their exposure to chemical synthesis, whereas the main focus of biologists is on understanding the interactions of native molecules. These divergent perspectives make collaborations a compelling strategy for the chemical biologist attempting to have an impact on biology because an engineering outlook can be grafted onto the specific expertise necessary to address questions at a biologically relevant level. An example of this approach is our collaboration with David Mangelsdorf and his laboratory on engineering the nuclear hormone retinoid X receptor (RXR) to have novel ligand specificity, studies in which we showed that the specificity for ligand-dependent transcriptional activation was surprisingly susceptible to alteration through mutagenesis [4].

We could not have done this project without access to the technical resources, materials, and background knowledge possessed by the Mangelsdorf laboratory. Conversely, the Mangelsdorf laboratory would have been unlikely to attempt a project aimed at engineering the specificity for ligand binding, creating a situation where both scientific outlooks were necessary to achieve results. Interestingly, soon after our paper was submitted for publication Gronemeyer and colleagues described some of the same mutations in the RXR. Rather than focusing on altering the specificity for ligand binding, Vivat *et al.* [5] characterized the high level of constitutive activation by one of the mutants, a result we had also observed but not emphasized, and used their observations to add details to the model of RXR-mediated transcriptional activation. Thus, similar mutagenesis experiments produced different lessons about the properties of RXR because of the different interests and experimental philosophies of the researchers involved.

Synergy of combining small-molecule synthesis and protein engineering

One of the defining characteristics of the chemical biology community is the ability to design small molecules with potentially useful properties and then synthesize the molecules needed to achieve the biological research aims. These capabilities are an advantage for chemistry-oriented laboratories because molecule design, synthesis and product analysis are difficult for laboratories that are oriented toward molecular biology. The advantage conveyed by expertise in small-molecule synthesis can also be exploited in protein engineering.

One strategy for using synthesis to amplify the potential of protein engineering is to design novel combinations of engineered proteins and synthetic ligands. One example of this approach from Schreiber and coworkers [6] is mutagenesis of cyclophilin A to enlarge its binding pocket and allow its recognition by a chemically modified cyclosporin that interacts poorly with wild-type cyclophilin A. The goal of this work is, eventually, to develop engineered ligand–receptor combinations that can be used to examine intracellular function in a more controlled fashion than is now possible.

Another example of this strategy is work by Shokat and coworkers [7,8] aimed at remodeling protein kinase specificity so that kinases use unnatural ATP nucleotide analogs rather than ATP for phosphoryl transfer. Because other kinases present in complex cell extracts cannot use the nucleotide analog, and the engineered kinase cannot readily use ATP, this approach offers the potential for specific labeling and identification of the natural substrates of the variant kinases, findings that would help elucidate details of signal transduction pathways. By themselves, neither the variant kinases nor the synthetic ATP analogs would be particularly interesting molecules. Combined, however, they create a new approach to understanding a significant biological problem. It is difficult to see how researchers who were not consciously working at the chemistry–biology interface could achieve such innovation.

Multi-step engineering: a new synthesis

In synthetic organic chemistry, products are built through multi-step procedures. Protein engineering, which can be viewed as synthetic chemistry for very large molecules, can mimic this approach by adopting difficult goals that require multiple steps to achieve dramatic transformations of structure or function. The multistep approach confers distinct advantages to protein engineers because it requires knowledge of the strengths and weaknesses of the spectrum of techniques available for protein engineering and a confidence that they can be applied successfully. Examples of multistep engineering include minimization of the size of atrial natriuretic peptide and protein A by Wells and colleagues [9,10] and the identification by Wrighton and coworkers [11] of a small peptide that is potent erythropoeitin mimic. In these studies the intermediate proteins and peptides, like most intermediates in organic synthesis, possessed few of the properties desired of the target molecule. By using repeated rounds of phage display or site-directed mutagenesis, however, the experimenters selected for favorable characteristics, such as reduced size or increased binding affinity. The systematic optimization resulted in a steady improvement in properties until eventually the target functions were achieved.

Clearly the analogy with synthetic chemistry can be taken too far as the palette of techniques available to the protein engineer is much smaller than the array of reactions available to the chemist, and our understanding of

the complexities of protein structure lags behind our understanding of small-molecule reactivity. As the rules for engineering proteins become better developed, however, the parallels will grow more striking and it is not out of the question that even the most sophisticated aspects of chemical synthesis, such as retrosynthetic analysis, might someday be applied to complex problems in protein engineering.

Renaissance of semi-synthetic enzymes?

Some of the earliest examples of protein engineering involved use of site-specific chemical modification of proteins to alter function [12–14]. The use of semi-synthesis in protein engineering has not been as prominent as the use of genetic manipulations because of the necessity to include a modification *in vitro* and because of limitations on our ability to selectively modify proteins. Recently, however, two groups [15–17] have introduced a strategy that could help reverse this trend by biologically expressing protein fragments with an intein [18,19]. The intein spontaneously rearranges, generating a thioester linkage in the mainchain that serves as a point of specific reactivity towards nucleophiles. To increase efficiency, the original polypeptide thioester linkage is exchanged through transthioesterification with a small-molecule thiol. Expressed proteins and synthetic peptides with an amino-terminal cysteine then attack this thioester, fusing the protein and peptide by first transthioesterification and then formation of a native peptide bond.

The method, termed expressed protein ligation [15], has been used to derivatize carboxy-terminal Src kinase with a phosphorylated peptide [15], to synthesize cytotoxic proteins RNAse A and *Hpa*I [16], and to modify the σ^{70} subunit of *Escherichia coli* RNA polymerase [17]. Perler and coworkers [19] have used similar methodology to express precursor protein fragments in separate hosts and then splice them *in vitro*.The value of expressed protein ligation relative to other methodologies for protein modification and cross-linking remains to be demonstrated, but its apparent efficiency and versatility suggest its potential to become a widely useful route to modification of protein termini by other proteins, peptides, oligonucleotides, and small molecules. Applications include the synthesis of modified proteins that are poorly expressed, the synthesis of proteins to contain unnatural amino acids or other non-native modifications, and the synthesis of semi-synthetic proteins that are too large to be obtained by chemical or enzymatic ligation of peptides.

Designing proteins for new applications

As noted above, one of the reasons to engineer proteins is to obtain therapeutics, experimental tools, or industrial reagents. Such variant proteins are difficult to obtain because combining altered function with adequate levels of activity is often challenging. To be successful, therefore, a thorough knowledge of the target protein needs to be combined with the selection of appropriate engineering strategies. One recent report that has important implications for development of improved enzymes for industry is the engineering of a thermolysin-like protease from *Bacillus stearothermophilus* to resist boiling by Van den Burg and colleagues [20]. Although the thermolysin was already moderately thermostable, introduction of a strategically placed disulfide combined with a series of 'rigidifying' mutations (such as glycine to alanine or varied amino acids to proline) capable of destabilizing a locally denatured state, increased the half-life at 100°C from being negligible to 170 minutes. The finding that highly thermostable proteins can be obtained by directed mutagenesis affords a new approach for engineering proteins to survive under harsh conditions, and supplies experimental evidence that some of the rules governing enzyme thermostabilty are beginning to be understood.

Summary

Protein engineering is used to manipulate biology in much the way that chemists are accustomed to manipulating chemical reactions. The studies cited here are informed by detailed knowledge of the biology under investigation and provide proof that the engineering perspective continues to afford important basic insights into macromolecular function. Success is no reason to become complacent, however. Protein engineering must always be put in a larger biological context, and the question of what protein engineering experiments are really teaching us must be rigorously and critically addressed.

Acknowledgements

We thank T.W. Muir and M-Q. Xu for providing us with their manuscripts prior to publication. This Crosstalk is dedicated to Elias J. Corey on the occasion of his seventieth birthday. D.R.C. is an Assistant Investigator with the Howard Hughes Medical Institute.

References

1. Tang, W-J. & Gilman, A.G. (1995). Construction of a soluble adenylyl cyclase activated by $G_s\alpha$ and forskolin. *Science* **268**, 1769-1772.
2. Tang, W-J., Stanzel, M. & Gilman, A.G. (1995). Truncation and alanine scanning mutants of type I adenylyl cyclase. *Biochemistry* **34**, 14563-14572.
3. Tesmer, J.J., Sunahara, R.K., Gilman, A.G. & Sprang, S.R. (1997). Crystal structure of the catalytic domains of adenylyl cyclase in a complex with $G_{s\alpha}$-GTPγS. *Science* **278**, 1907-1916.
4. Peet, D.J., Doyle, D.F., Corey, D.R. & Mangelsdorf, D.J. (1998). Engineering novel specificities for ligand-activated transcription in the nuclear hormone receptor RXR. *Chem. Biol.* **5**, 13-21.
5. Vivat, V., Zechel, C., Wurtz, J-M., Bourguet, W., Kagechika, H., Umemiya, H., Shudo, K., Moras, D., Gronemeyer, H. & Chambon, P. (1997). A mutation mimicking ligand-induced conformational change yields a constitutive RXR that senses allosteric effects in neterodimers. *EMBO J.* **16**, 5697-5709.
6. Belshaw, P.J., Schoepfer, J.G., Liu, K-Q., Morrison, K.L. & Schreiber, S.L. (1995). Rational design of orthogonal receptor–ligand combinations. *Angew. Chem. Int. Ed. Engl.* **34**, 2129-2132.
7. Shah, K., Deirmengian, C. & Shokat, K.M. (1997). Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc. Natl Acad. Sci. USA* **94**, 3565-3570.
8. Liu, Y., Shah, K., Yang, F., Witucki, L. & Shokat, K.M. (1998). Engineering Src family protein kinases with unnatural nucleotide specificity. *Chem. Biol.* **5**, 91-100.

9. Li, B., Tom, J.Y.K., Oare, D., Yen, R., Fairbrother, W. J., Wells, J.A. & Cunningham, B.C. (1995). Minimization of a polypeptide hormone. *Science* **270**, 1657-1660.
10. Braisted, A.C. & Wells, J.A. (1996). Minimizing a binding domain from protein A. *Proc. Natl Acad. Sci. USA* **93**, 5688-5692.
11. Wrighton, N.C., Farrell, F.X., Chang, R., Kashyap, A.K., Barbone, F.P., Mulcahy, L.S., Johnson, D.L., Barrett, R.W., Jolliffe, L.K. & Dower, W.J. (1996). Small peptides as potent mimetics of the protein hormone erythropoeitin. *Science* **273**, 458-463.
12. Neet, K.E. & Koshland, D.E. (1966). The conversion of serine at the active site of subtilisin to cysteine: a "chemical mutation". *Proc. Natl Acad. Sci. USA* **56**, 1606-1611.
13. Polgar, L. & Bender, M.L. (1966). A new enzyme containing a synthetically formed active site. Thiol-subtilisin. *J. Am. Chem. Soc.* **88**, 3153-3154.
14. Kaiser, E.T. & Lawrence, D.S. (1984). Chemical mutation of enzyme active sites. *Science* **226**, 505-511.
15. Muir, T.W., Sondhi, D. & Cole, P.A. (1998). Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. USA* **95**, 6705-6710.
16. Evans, T.C., Jr., Benner, J. & Xu, M.-Q. (1998). Synthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.*, in press.
17. Severinov, K. & Muir, T.W. (1998). Expressed protein ligation: a novel method for studying protein-protein interactions in transcription. *J. Biol. Chem.*, in press.
18. Perler, F.B., Xu, M.-Q. & Paulus, H. (1997). Protein splicing and autoproteolysis mechanisms. *Curr. Opin. Chem. Biol.* **1**. 292-299.
19. Southworth, M.W., Adam, E., Panne, D, Byer, R. Kautz, R. & Perler, F.B. (1998). Control of protein splicing by intein fragment reassembly. *EMBO J.* **17**, 918-926..
20. Van den Burg, B., Vriend, G., Veltman, O.R., Venema, G. & Eijsink, V.G.H. (1998). Engineering an enzyme to resist boiling. *Proc. Natl Acad. Sci. USA* **95**, 2056-2060.