PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Genetically Engineered Reagents [and Discussion]

T. Atkinson, R. F. Sherwood, R. F. C. Mantoura, W. J. Albery, M. Cresser, J. D. R. Thomas and G. S. Wilson

Phil. Trans. R. Soc. Lond. A 1990 **333**, 93-98
doi: 10.1098/rsta.1990.0140

Genetically engineered reagents

By T. Atkinson and R. F. Sherwood

Division of Biotechnology, PHLS Centre for Applied Microbiology and Research, Porton, Salisbury SP4 0JG, U.K.

The development of recombinant DNA technology has led to the wide availability of reagents for the analysis of trace quantities of macromolecules and small chemical agents. The cloning and expression of genes coding for the synthesis of proteins and enzymes has meant that previously scarce protein molecules have now become available as both standards for their own measurement in biological samples and as reagents for the determination and measurement of other molecules. In addition the advent of recombinant DNA technology has led to the development and advancement of valuable 'spin off' capabilities such as (i) site directed mutagenesis providing the ability to specifically alter the amino acid sequence and structure of protein molecules at the level of the gene, (ii) the synthesis of DNA probes to provide for both the isolation and detection of specific gene sequences in biological samples and (iii) gene amplification employing techniques such as the polymerase chain reaction to amplify and provide multiple hundred to thousands of copies of the specific gene sequence.

The ability to clone genes from a wide variety of cell sources and then promote their enhanced expression in a system of choice has led to the production of large quantities of specific proteins not previously available to the scientist (Atkinson *et al.* 1986; Sherwood *et al.* 1990). The variety of host expression systems available and their specific advantages and typical yields are illustrated in table 1. The range of previously unavailable proteins which have now become available for use in the analysis or diagnosis of analytes in body fluids or for environmental analysis includes antibiotic degrading, or modifying enzymes, and enzymes for lipid and triglyceride analysis. In addition many molecules such as the immune modulators interferon and interleukin, which were previously relatively unavailable have now become available in quantity, as a result of cloning and expression, for use as therapeutic agents as well as standards for the determination of such molecules in biological samples. Similarly the limited availability of enzymes from their natural host, such as firefly luciferase, used both in the determination of ATP and for signal amplification, has been removed by cloning the gene and obtaining expression in several different host expression systems. The marked variation in the quality of some protein and particularly enzyme reagents when normally isolated from their natural host has also been circumvented by recombinant DNA technology. This has been demonstrated by the chemical synthesis of a gene coding for the synthesis of horse radish peroxidase, an enzyme widely used as a detector system in clinical diagnosis, and its expression in a microbial host, removing the need to isolate the protein from plant tissue of an inherently variable nature. Other large peptides or protein molecules such as hormones and toxins have also been made available through gene cloning and expression. Examples of such proteins are presented in table 2.

Phil. Trans. R. Soc. Lond. A (1990) **333**, 93–98
Printed in the Great Britain

Table 1. *Levels of expression of proteins and peptides when the gene encoding the protein or peptide is cloned in a particular host*

host	level of expression in mg per l culture	advantages/disadvantages
E. coli	250–3000	(*a*) some products formed as insoluble inclusion bodies requiring renaturation (*b*) no glycosylation of mammalian proteins (*c*) rapid and easy growth conditions
B. subtilis	50–3000	(*a*) secretion into external media can be sometimes obtained (*b*) no glycosylation of mammalian proteins (*c*) high level expression more difficult to achieve than *E. coli*
Streptomyces	50–2500	(*a*) secretion into external media can be sometimes obtained (*b*) some form of glycosylation may occur (*c*) protracted growth often required
yeast	(i) 50–1000 for proteins (ii) 5–50 for peptides	(*a*) secretion into external media is excellent for peptides (*b*) some form of glycosylation may occur (*c*) protracted growth often required
mammalian cells	50–200	(*a*) secretion into external media can often be arranged for products normally secreted (*b*) glycosylation occurs as per normal protein for mammalian proteins (*c*) amplification sometimes difficult, growth protracted
insect cells	50–400	(*a*) similar to mammalian cell systems (*b*) glycosylation is identical to mammalian cells (*c*) system still in development

Table 2. *Examples of proteins expressed by recombinant DNA techniques*

(1) hormones, immune modulators and receptor active proteins	interferons (α, β and γ), interleukins, growth hormones tumour necrosis factor granulocyte macrophage colony stimulating factor
(2) viral antigens	HIV I (AIDS), hepatitis, herpes simplex
(3) marker enzymes	β-galactosidase, alkaline phosphatase, malate dehydrogenase horse radish peroxidase, firefly and bacterial luciferases
(4) other proteins of diagnostic value	protein A, protein G, streptavidin, mussel adhesive protein

Genetic engineering has also been used for the construction of novel chimaeric hybrid fusion proteins with both a specific binding domain and an enzymic activity. *Staphylococcal* Protein A is an immunoglobulin binding protein which binds IgG subclasses 1, 2 and 4 and under appropriate conditions also IgA. Similarly,

Streptococcal Protein G binds all subclasses of IgG, while *Peptostreptococcal* Protein L binds the light chain of all human immunoglobulins. Cloning and subsequent genetic isolation of the repetitive IgG binding domains of Protein A and Protein G has allowed their genetic fusion to cloned marker or 'detector' enzymes such as β-galactosidase or alkaline phosphatase to produce a gene coding for the synthesis of a single protein with both IgG binding and enzyme activity. Such proteins are of value as universal second antibody reagents in immunoassay. Similarly, the genetic fusion of the well-determined cellulose binding domain or the mussel 'adhesive' protein provides a means of anchoring a specific enzyme to a solid matrix for use as a reagent. The biotin binding protein streptavidin from *Streptomyces* with a dissociation constant for biotin of about 10^{-15} has one of the strongest affinities for its substrate of any known protein (Wilchek & Bayer 1989). Genetic hybrids of the biotin binding domain of streptavidin to enzyme markers provides a means for binding such chimaeric fusion proteins to biotinized or biotinylated proteins or solid supports, for which a plethora of 'activated' biotin analogues already exists.

The past decade has seen the realization of our ability to synthesize oligonucleotide sequences which code for part, or indeed all, of the information necessary to result in the expression of a specific protein. In addition we have developed the capability of using such oligonucleotides to mispair with the native gene and thus introduce, site-specifically, mutations at predetermined points in the gene coding for the amino acid sequence of a protein. It has therefore become possible to 'engineer' enzymes to new activities and specificities or to otherwise beneficially modify enzymes for their intended use (Shaw 1987). Early work in this field concentrated on modification to enzymes for which a detailed X-ray derived crystallographic structure existed, thus allowing the effect of an intended change to be predicted from consideration of the amino acid interactions in the molecule. With the development, however, of high resolution molecular graphics and enhanced computer-based structure prediction (particularly of a protein whose structure is itself not known but which is one of a related family of proteins of which one member has been structurally determined by X-ray crystallography) the field of work has expanded to include educated experimentation on the introduction of desired modifications into proteins whose structure has not been well determined or is only poorly known.

The redesign of enzymes is one of the major areas of enzyme engineering and many changes are possible including K_m and k_{cat}, pH optimum, removal of effector sites, if present, change of co-factor requirement in the case of dehydrogenase, thermal stability, etc. In the case of lactate dehydrogenase from *B. stearothermophilus*, an enzyme whose structure has been modelled on the structure of lactate dehydrogenase crystallized from Dogfish, the majority of the above parameters have been changed (Clarke *et al.* 1989*a*, *b*). Perhaps the most startling direct achievements of this work have been (*a*) the removal of the fructose 1,6 bis-phosphate binding site by changing the Arg at residue 173 to Gln (Arg 173 Gln), a change which removes the intersubunit charge repulsion and results in a 50-fold increase in the enzyme's affinity for pyruvate, (*b*) the 4.5-fold increase in thermal stability of this already highly thermostable protein ($t_{\frac{1}{2}}$ 90 °C from 6 to 27 min) by Ile 250 Asn which removes the problem hydrophobic residue from a hydrophilic pocket from which 'melting' of the protein is first initiated and (*c*) the quite deliberate change of this lactate dehydrogenase into a malate dehydrogenase accepting the extra carbon atom and carboxyl group on the new substrate by Gln 102 Arg resulting in a seven orders of magnitude change in k_{cat}/K_m and a malate dehydrogenase with a k_{cat}/K_m for the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

new substrate oxaloacetate equal to that of the original wild type enzyme against pyruvate. In addition the new protein turned over oxaloacetate or malate twice as rapidly as the native wild type malate dehydrogenase isolated from *B. stearothermophilus*. Recent work by the same group on both the original lactate dehydrogenase and also malate dehydrogenase from *E. coli* have shown how to engineer pH optima changes, alterations in co-factor use and have developed methods for 'reporting' on the mechanism of protein folding or unfolding by studying the time-resolved fluorescence of Trp residues engineered into the molecule at various points in the structure of a protein from which all Trp residues have been deleted and replaced by Tyr. This has the added advantage that the protein reagent is genetically marked and the protein recognizable in a range of similar proteins. The eventual aim of such work is to be able to specifically redesign enzymes and proteins to have the thermal, chemical and protease stability and specificity desired for a particular analysis.

Monoclonal antibodies (MABS), developed with particular specificities, have found use in recent years as highly sensitive and specific binding agents (Winter 1989). However, it is possible to conceive of MABS as molecules with 'enzyme potential' and hence Abzymes (MABS with catalytic activity) have been designed and isolated, particularly in respect of transition state intermediate chemicals (Lerner & Tramontano 1987). The maximum enhancement of non-enzymic conversion of a substrate that has been obtained with a specific Abzyme is about 6×10^6 fold. Abzymes active in the hydrolysis of ester, amide and carbonate derivatives of aliphatic and aromatic compounds have been obtained, as have 'peptidase' (peptide hydrolysing) Abzymes. The dissociation constant of any Abzyme for its substrate is usually about 10^{-12} and the K_m for enzymic activity is usually in the range micro- to nanomolar, a two to three order of magnitude better affinity constant than conventional enzymes. A k_{cat} as high as 20 has been obtained with Abzymes demonstrating that some MABS are now well advanced to becoming highly respectable enzymes (Green & Tawfik 1989). Since the antigenic binding area of a MAB is small compared with the overall size of the molecule it is possible to think of engineering additional and different antigenic binding sites into MABS or to design in specific effector sites, e.g. the mollusc 'adhesive' protein decameric serine and tyrosine rich repeat which can be used to induce binding to solid matrices or surfaces, cell surface receptors or thiol rich domains for chemical coupling. Certainly such molecules have remarkable potential for the design of novel analytical systems.

Recombinant DNA technology has also been responsible for the development of DNA probes for the diagnosis of bacterial or viral contamination of body fluids and food or environmental samples (Viscidi & Yolken 1987; Matthews & Kricka 1988). In contrast to the predominantly small probes used for the isolation or cloning of specific genes, DNA probes used in detection or identification of a particular organism tend to be large to attain the degree of discrimination required between species which may be similar. Such probes can be used for the detection of toxin producing organisms which may themselves either be pathogenic or indeed non-pathogenic, for the specific detection of pathogenic organisms or virus or in forensic medicine for the specific detection and identification of a human chromosome in 'genetic fingerprinting'. The detection systems used for the measurement of specific probe to chromosome hybridizations may be radioactive or protein mediated. One such system is based on the binding of streptavidin-enzyme conjugates to biotinylated probes although other systems have been designed.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The ability to design probes capable of identifying specific DNA sequences, the ability to redesign the binding and catalysis of proteins and to develop the overproduction of molecules synthesized by complex metabolic pathways will lead to further significant developments in trace analysis in the future.

References

Atkinson, T., Barstow, D. A., Court, J., Minton, N. P., Sharp, R. J. & Sherwood, R. F. 1986 High-level microbial expression and purification of recombinant proteins. In *Bioactive microbial products 3* (ed. J. D. Stowell, P. J. Bailey & D. J. Winstanley), vol. 3, pp. 27–44. London: Academic Press.

Clarke, A. R., Atkinson, T. & Holbrook, J. J. 1989*a* From analysis to synthesis: new ligand binding sites on the lactate dehydrogenase framework. Part I. *Trends biochem. Sci.* **14**, 101–105.

Clarke, A. R., Atkinson, T. & Holbrook, J. J. 1989*b* From analysis to synthesis: new ligand binding sites on the lactate dehydrogenase framework. Part II. *Trends biochem. Sci.* **14**, 145–148.

Green, B. S. & Tawfik, D. S. 1989 Catalytic monoclonal antibodies: tailor-made-enzyme-like catalysts for chemical reactions. *Trends Biotechnol.* **7**, 304–310.

Lerner, R. A. & Tramontano, A. 1987 Antibodies as enzymes. *Trends biochem. Sci.* **12**, 427–430.

Matthews, J. A. & Kricka, L. J. 1988 Analytical strategies for the use of DNA probes. *Analyt. Biochem.* **169**, 1–25.

Shaw, V. 1987 Protein Engineering. The design, synthesis and characterization of factitious proteins. *Biochem. J.* **246**, 1–17.

Sherwood, R. F., Plank, R., Baker, J. & Atkinson, T. 1990 Large scale production of proteins from recombinant DNA in *E. coli*. In *Recombinant DNA methods in microbiology* (ed. J. M. Grange), Society for Applied Bacteriology Technical Series, vol. 28. (In the press.)

Viscidi, R. P. & Yolken, R. G. 1987 Molecular diagnosis of infectious disease by nucleic acid hybridisation. *Molecular Cellular Probes* **1**, 3–14.

Wilchek, M. & Bayer, E. A. 1989 Avidin-biotin technology ten years on: has it lived up to its expectations? *Trends biochem. Sci.* **14**, 408–412.

Winter, G. P. 1989 Antibody engineering. *Phil. Trans. R. Soc. Lond.* B **324**, 537–547.

Discussion

R. F. C. Mantoura (*Plymouth Marine Laboratory, U.K.*). What are the possibilities of escape of genetically engineered organisms and how would they affect the environment and water pollution?

T. Atkinson. The growth of production of genetically engineered organisms is tightly controlled and regulated. Environmental release is also tightly regulated and is the subject of current legislation. In general terms the finite chance of the release of something undesirable is negligible. Laboratory experiments are also being carried out in the U.K. on the effects of such a release by researching laboratory based experimentally simulated eco-systems.

W. J. Albery. That sounds very reassuring. If its nasty it won't get out, and if it gets out its not nasty.

M. Cresser (*Aberdeen University, U.K.*). The point about possible contamination of water by genetically engineered organisms is an interesting one. At Aberdeen we are doing research on the factors regulating the migration of micro-organisms through soils for the U.K. Department of the Environment. Genetic engineering has provided

a unique opportunity in that we can use organisms tagged with a luminescent gene as an investigatory tool. To tackle the problem in any other way is virtually impossible. Thus it is rather a chicken-and-egg situation. However, it does mean that engineering techniques will facilitate a better risk assessment of the use of engineered materials.

W. J. Albery (*Molecular Sensors Unit, University of Oxford, U.K.*). Protein engineering sounds wonderful, but I would like to ask Professor Atkinson what chances does he have of getting it right when he designs the changes to his molecule? Does he get it right every time, one in ten, one in a hundred, etc.?

T. Atkinson. The chances of getting it right are proportional to our pre-existing state of knowledge of the molecule. The more we know, the better the three-dimensional structure then the greater the chance. At best the chances are certainly better than ten to one.

J. D. R. Thomas (*School of Chemistry and Applied Chemistry, University of Wales, U.K.*). If genetic engineering shifts the optimum pH of enzymes by 1.5 units, would the effect of immobilization (which also shifts the optimum pH) be cumulative to provide an even bigger overall shift in optimum pH, or would there be some other influence?

T. Atkinson. It solely depends on the reasons for the shift. In general terms it is difficult to shift a pH optimum by more than 1.5 pH units. Immobilization may be additive to the protein engineered shift, it may detract from it; it would purely depend on the different interactions and how they altered structure and kinetics.

G. S. Wilson (*University of Kansas, U.S.A.*). What are the possibilities of protein (antibody, enzyme) engineering when the X-ray structure is not known? What about structural prediction based on amino acid sequences?

T. Atkinson. Structural prediction can be achieved based on amino acid sequence, but, at best, is not more than 70% accurate. Therefore protein engineering in the absence of an experimentally determined structure can be done but is very unpredictable. If a three-dimensional structure does not exist *per se*, the next best thing is where one of the family of that particular protein has had its three-dimensional structure determined. It is then possible to do molecular modelling to fit your amino acid sequence. In fact all the work carried out on the *B. stearothermophilus* lactate dehydrogenase that I have described was achieved using such a structure modelled on the same enzyme from Dogfish. Structural prediction from amino acid sequence is improving but is not yet accurate enough to give the precise detail necessary to give good 'odds' in protein engineering.