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

Commercial Prospects for Enzyme Engineering [and Discussion]

B. S. Hartley, N. L. Brown, R. W. Carrell, Mary Waye and M. F. Perutz

Phil. Trans. R. Soc. Lond. A 1986 **317**, 321-331
doi: 10.1098/rsta.1986.0042

Phil. Trans. R. Soc. Lond. A **317**, 321–331 (1986)
Printed in Great Britain

Commercial prospects for enzyme engineering

By B. S. Hartley, F.R.S.
Centre for Biotechnology, Imperial College, London SW7 2AZ, U.K.

Despite great enthusiasm for enzyme technology over the last two decades, the number and scale of commercial applications remains disappointing. To change this, protein engineering must concentrate on improving the properties of current commercial enzymes for specific process needs. Construction of novel enzymes with robust properties and tailored specificities will create new markets. The technology to do this is already fast and efficient, but lack of relevant tertiary-structure models is a serious constraint.

Improved thermostability could be achieved by introducing new sulphide bridges, increasing internal bonding and modifying surface charges. There are other obvious modifications that should improve stability to oxidation, extremes of pH, and heavy metals. These principles are illustrated by work in progress on subtilisin which is the major industrial enzyme.

Industrial enzymes

Because I am not involved in the production or sales of any enzyme I have little direct knowledge of the commercial processes in which new enzymes might be useful. I am an enthusiast for enzymes in general, and for enzyme engineering in particular, so might be accused of over-optimism. However, I have had sufficient experience in mixing with industrial colleagues over the last few years to understand their constraints and their apparent reservations about the commercial impact of this new technology, and will try to explain this.

Table 1 shows estimates of world markets for industrial enzymes, taken from the sources indicated. For the benefit of most academics present I should explain that such market information is generally only an approximation. For example, a more recent survey by the U.S. Department of Commerce (1984) estimates the current world market for industrial enzymes at 22 M lb per annum, at prices ranging from U.S. \$20–700 per pound, giving a market value around U.S. \$1 billion per annum (because the bulk enzymes are the cheapest). Such figures sound attractive until one asks what are the research and development costs, production costs, market volume, and extent of competition for any one particular enzyme. In sheer volume, the proteases used for detergents and food processing and the amylases for starch hydrolysis outweigh all others, so one must ask why this is. The answer is, I believe, that they have been around for a long time and industry and consumers have got used to them. Moreover, table 1 shows that the current bulk markets for enzymes are predominantly in food processing and brewing. These are very conservative industries because the regulatory constraints in introducing a new product or even a process improvement are severe. At best, this causes a delay which ties up the capital already expended, and at worst, can eliminate the perceived profit.

Another factor that influences the market for industrial enzymes is that the catalyst is generally only a small fraction of the cost of the final product. This point is obvious from comparing the prices of the enzymes shown in table 1. These do not reflect the production costs

Table 1. Major industrial enzymes

	U.S.$ kg^{-1} (1980)[a]	market size		
		tonnes	U.S.$m (1980)[a]	U.S.$m (1983)[b]
bacterial proteases	124	530	66	135
glycoamylase	102	350	36	?
α-amylase	38	320	12	20
glucose isomerase	800	70	56	70
rennin (chymosin)	2600	26	64	79
total	—	—	234	300

[a] Eveleigh (1981).
[b] Kidder & Peabody (1984) personal communication.

nor the actual specific activities in what is generally a very crude product, but they do rely on the unique properties of that particular enzyme. Thus chymosin, the protease in rennet, continues to command a high price for cheese-making even though other acid proteases will clot milk as efficiently; the side specifities of such other enzymes are considered bad for flavour and texture. The comparison of glucose isomerase and amylases is also revealing. The high price of the former reflects its unique properties of long-term stability and catalytic efficiency as an immobilized enzyme in a continuous process at relatively high temperatures.

I have concentrated entirely on major industrial enzymes because I believe that it is here that the considerable research and development costs involved in protein engineering will first be commercially justified. There is a large, rapidly growing, market for diagnostic enzymes, particularly with the advent of biosensors, but the overall market for any one of these enzymes limits the amount of research and development that can be afforded. However, this will not necessarily apply to novel enzymes, which are designed to allow analysis of a specific substance at high sensitivity. If the market for a particular biosensor were large, and no known natural enzyme could compete, there might be a sound case for some intelligent protein engineering.

The greatest long-term prospect for protein engineering lies in creating new markets for enzymes in processes and products in which they are not currently used. The ingenuity of organic chemists in creating a huge petrochemical industry is our paradigm. All agree that some day biomass will have to replace oil as a source of bulk organic chemicals. Enzymes are the ideal tools to manipulate this raw material, because they can perform sophisticated chemical transformations in the most complex mixtures. As enzyme engineers we are confident that we will be able to extend the range of their specifities and thereby increase the range of potential products. However, we must become the allies and not the competitors of the synthetic organic chemist, because there will always be reactions that he can do more effectively, just as we will often win when stereospecific substitutions are required.

Protein engineering technology

The moral for enzyme engineers must therefore be to design enzymes which are ideal for a specific process. However, before considering the various targets that might be achievable it is appropriate to survey the technology that lies behind protein engineering.

Figure 1 summarizes the strategies. Given as little as 1 mg of a pure protein it is now possible

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

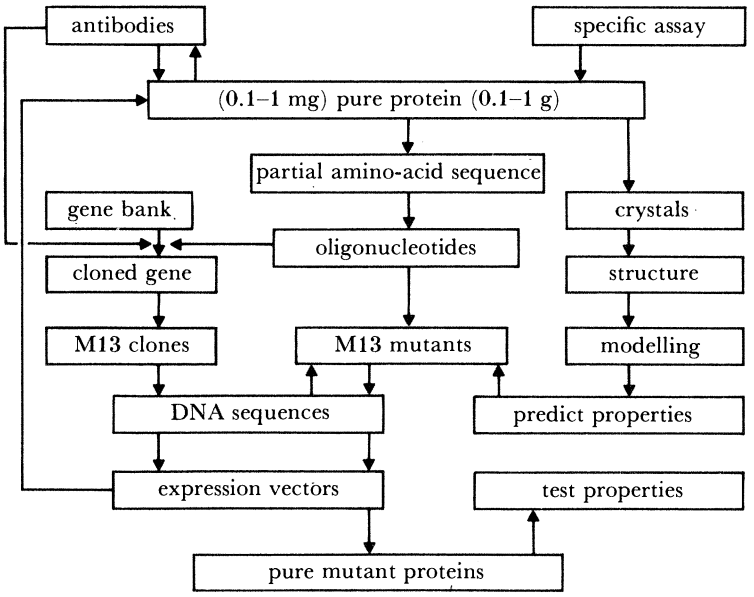

FIGURE 1. The strategies of protein engineering.

to clone the gene or the cDNA (c is complementary) that encodes it. The protein can be used to raise monospecific antibodies (preferably polyclonal for this purpose) which can be used to screen a gene bank or cDNA bank in which the inserts are cloned in front of a microbial promoter so that there is a reasonable chance that it will be expressed. Alternatively, and more generally, degenerate oligonucleotide probes can be synthesized if partial amino-acid sequences can be obtained, particularly around tryptophan and methionine residues which have low codon redundancy. The newer, more sensitive, sequencing techniques involving gas-phase sequenators or mass spectrometry are of particular value here. It is now almost certain that any gene can be rapidly cloned via such synthetic oligonucleotide probes.

From the cloned gene to the complete DNA sequence via subcloning into various single-stranded M13 phage vectors is now a matter of weeks, so the complete amino-acid sequence becomes quickly available. Manipulation of the upstream signal sequences into various expression vectors then results in a high probability that quantities of the protein can be obtained from fermentations of transformants of *E. coli*, yeast or *B. subtilis*. Moreover, as we have seen from previous talks, the technology of changing any codon for any other in the M13 clones is now extremely rapid and efficient. There is also a good chance that the mutant genes will express in the same expression vector as was used for the wild-type gene, so at this level the technology of protein engineering is efficient and rapid.

But we now hit the major stumbling block in exploitation of this new technology. Protein variants can be produced mindlessly with these tools, but the problem of purifying these to test their properties is relatively tedious and is justified only when reason suggests that the result will be significant. In almost all cases intelligent proposals for modifying amino-acid sequences will be based on a model of the tertiary structure, and this can only arise today via X-ray crystallography. Despite the improvements in data collection and handling made since the pioneering days of Kendrew and Perutz this remains, alas, the rate-limiting step.

The major constraint is generally the lack of suitable crystals. Protein crystallization is still

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

more of an art than a science and there has been disappointingly little effort and ingenuity in overcoming this fundamental hurdle. I do not entirely blame X-ray crystallographers for this because there are a host of proteins worthy of study that do crystallize and the research grant system does not encourage speculative novel solutions to methodological problems. Nevertheless, the wide commercial exploitation of protein engineering demands tertiary structures for particular proteins that have not previously been studied, and we cannot continue to rely on art and luck to get good crystals.

Given suitable crystals, a tertiary structure should emerge, at least at the 3 Å† level. This may be adequate for protein engineers because they themselves can make mutants to help verify the model building. But there is an international shortage of experienced protein crystallographers; there is also a shortage of protein chemists, who are in equally great demand to purify proteins, to partly sequence them, test physical and enzymic properties, and perform intelligent chemical modifications.

The area of modelling mutants from a known structure has been revolutionized by the latest tools of molecular graphics. Although this is a key element in the whole technology and has attracted much interest (for example, the recent E.E.C. 'Bioinformatics' programme), it is not a rate-limiting step. An intelligent protein chemist or crystallographer can come up with a dozen interesting prospective mutants in a single day's dialogue with a well programmed computer. However we all hope that collaboration between experimental and theoretical protein engineers will eventually come up with predictive algorithms for protein structure, stability and properties. I merely stress that the urgent immediate needs are for more experimental data on which to base such theories.

Commercial targets for protein engineers

In the long run I feel certain that the prediction of protein structure and properties from gene sequences will prove to be the most important target for all of us, but there are some more immediate opportunities that could be grasped. Returning to the more mundane field of industrial enzymes, it is obvious that many of these could be improved and that such improvements would bring quick returns to those who get there first with the best. Table 2 lists some of the targets and some possible solutions.

I have already mentioned that long-term stability is often more important than high catalytic activity for many processes. High temperatures, oxidizing agents, heavy metals or organic solvents are the major process hazards; what can we do about these?

There is no doubt that disulphide bridges contribute greatly to the stability of many extracellular proteins. This can be readily seen from the fact that the reduced proteins always show the physico-chemical characteristics of 'random coil' under all conditions. There can be little steric problem in accommodating two thiol groups in place of a disulphide and thiol groups are found in hydrophobic regions in many proteins, so these factors are hardly adequate to explain the instability of the reduced structure. The great difference in free energy of a covalent disulphide bond versus the pair of hydrogen bonds that would replace them is surely the obvious explanation, but this conclusion continues to be resisted in many text books and by many biochemists.

Perhaps their scepticism arises from the fact that all intracellular proteins and some

† $1\ \text{Å} = 10^{-1}\ \text{nm} = 10^{-10}\ \text{m}$.

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

Table 2. Protein engineering targets

(1) improved thermostability
  introduce internal or surface disulphide bridges
  increase internal hydrogen bonds
  improve internal hydrophobic packing
  increase surface salt bridges
(2) stability to oxidation
  convert Cys to Ala or Ser
  convert Met to Gln, Val, Ile or Leu
  convert Trp to Phe or Tyr
(3) stability to heavy metals
  convert Cys to Ala or Ser
  convert Met to Gln, Val, Ile or Leu
  alter surface carboxyls
(4) pH stability
  alter surface charged groups
  replace internal His, Cys and Tyr
  replace internal ion pairs
(5) improved enzymic properties
  altered specificity
  increased turnover number
  altered pH profile

extracellular ones are stable in the absence of disulphide bridges. But this ignores the fact that proteins are synthesized in the glycolytic compartment of the cell, which is highly reducing because of the redox potential of $NADH/NAD^+$; hence a disulphide bridge would be thermodynamically unstable in this environment. Once a protein is secreted across the cell membrane (or endoplasmic reticulum in eukaryotes) it is in equilibrium with molecular oxygen, so thiol groups become unstable and disulphide bridges are favoured. These chemical arguments are supported by the biological evidence that almost no proteins in the glycolytic compartment have disulphide bridges (except where they form part of a catalytic site as in lipoate dehydrogenase). Some extracellular proteins have free thiol groups, generally at active sites (for example, papain, which exists naturally as an inactive disulphide form), but such proteins are unstable in the presence of air. Indeed one should avoid the assumption that all proteins are designed by nature for optimum stability; a limited lifetime may be an evolutionary advantage. However, the protein engineer must design his enzymes for the harsh world of industrial processes.

One would think, therefore, that new disulphide bridges should increase protein thermostability. Internally they are often found to connect domains of secondary structure that would otherwise be floppy. In engineering new bridges one must be careful to avoid steric distortions that would dislocate such domains. One rule of thumb is to search by computer for pairs of contiguous side chains that are effectively isosteric with a disulphide bridge.

Figure 2 shows that appropriate internal locations can often be found in a pair of serine residues which hydrogen bond to each other. More rarely, the methyl group of a methionine which is in contact with the $-CH_2-$ of a main-chain glycine residue meets the criterion. A disulphide bridge replacing a Thr···Ser interaction would leave a hole in place of the threonine methyl group; this might be mended by converting an adjacent Val to Leu or Ile. Similarly a replacement of Val···Val by Cys–Cys would leave two 'methyl holes' to be filled.

Such steric constraints would not apply to new surface disulphide bridges connecting, for example, two surface Ala residues as shown in figure 2. There are many locations in which

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

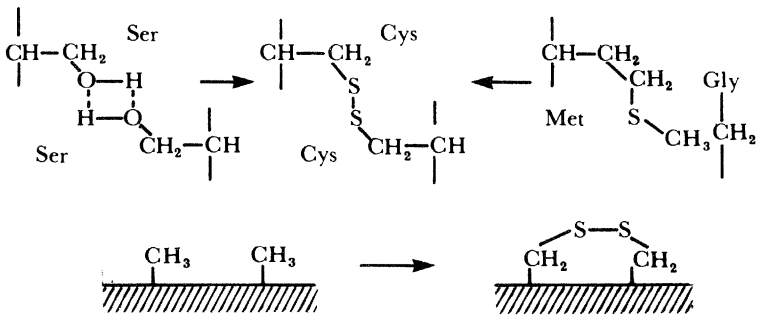

FIGURE 2. Introduction of new disulphide bridges: (*a*) shows that internal isosteric substitutions might be made wherever two internal Ser residues are hydrogen bonded to each other or wherever the methyl group of Met is in contact with the H atom attached to the α-C of Gly; (*b*) indicates that surface bridges can be grafted in place of a pair of surface Ala residues with appropriate distance and angles.

these might lock adjacent domains of secondary structure. Such surface bridges would eliminate problems connected with the kinetics and thermodynamics of refolding the mutant protein and oxidizing the two Cys residues after it has been synthesized (and possibly secreted) by the host organism. A possible commercial disadvantage is that these new surface bridges might be prone to oxidation or to interaction with heavy metal ions.

A recent example which validates these concepts arises from the work of Perry & Wetzel (1984), who replaced Ile-3 with Cys in T4 phage lysozyme. The resulting enzyme had thiol groups at Cys-3, Cys-54 and Cys-97 but, as expected from molecular graphics modelling, the Cys-3–Cys-97 bond formed preferentially on oxidation. The half life at 67 °C increased from 11 min to over 6 hr provided that the mutant enzyme was also blocked with iodoacetate to avoid oxidation problems. This demonstrates the power of protein engineering to improve both thermostability and stability to oxidation.

We lack good theories for protein thermostability, but maximizing internal hydrogen bonding is one possibility. Another is to maximize space filling by the internal hydrophobic residues lying between structural domains. Perutz & Raidt (1975) have also drawn attention to the increased thermostability of a bacterial ferredoxin in which two surface ion pairs (Lys···Glu and His···Glu) replace the sterically similar Lys···Gln and His···Gln of the less stable species. All of these avenues can be explored by protein engineers.

Cysteine, methionine and tryptophan residues are particularly sensitive to oxidation. Heavy metal ions bind to Cys and Met residues and to clusters of carboxyl groups. If these residues lie at the surface, the effects of oxidation or liganding on stability may be minimal, but because they are usually buried the consequence of modification will be irreversible denaturation. Once again substitutions that cause minimal steric distortions could be modelled.

Similar principles apply to buried groups that may ionize at extremes of pH such as His, Cys, or Tyr, leading to pH denaturation. Internal ion pairs involving Arg, Lys, His, Asp, or Glu might be replaced by neutral residues which form a hydrogen bond. However, surface charges must surely contribute to pH stability by their dielectric effects across the structure and this phenomenon needs to be explored.

Finally, the most important class of substitutions is those that alter the catalytic properties of the molecule. I will not discuss these here because each enzyme will present its own possibilities. Modifications of substrate specificity or of pH activity profiles may have major

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

industrial potential. I must also emphasize the observation made by Professor Fersht and his colleagues (this symposium), that a massive increase in k_{cat}/K_M for tyrosyl-tRNA synthetase could be achieved by intelligent substitutions in the catalytic site (Wilkinson *et al.* 1984). It allows the possibility that protein engineers can improve on Nature, at least for the narrow purposes for which they wish to use an enzyme.

Protein engineering of subtilisin

This topic will be covered in greater depth by Dr Wells (this symposium), but I use it here to illustrate the general points made above. Professor Fersht and I also chose this enzyme as a test bed for commercially relevant protein engineering for the following reasons.

(*a*) It is a serine protease of known structure (Wright *et al.* 1969; Drenth *et al.* 1972) and mechanism (Blow *et al.* 1969; Robertus *et al.* 1972), which is already stable at high temperatures without the benefit of disulphide bridges.

(*b*) A similar enzyme is found in a wide range of bacilli, which secrete it in huge amounts into the medium in a 'pre-sporulation burst' when the cells are starved. These many species of subtilisins already form a library of enzymes that have been 'protein engineered' by Nature.

(*c*) It is already the major industrial enzyme through its use in enzyme detergents. Hence even small improvements would have early value.

(*d*) For that reason industry has undertaken a lot of conventional screening and strain improvement which results in superproducing strains that secrete more than 2 g of enzyme per litre of fermentation broth. This greatly eases the task of purifying large quantities of mutant enzymes.

(*e*) Genetic engineering of *B. subtilis* and allied strains is well developed so there is good prospect of replacing an improved gene directly into the current industrial organism.

Since October 1983 we have made gene banks of *B. amyloliquefaciens* in plasmid pBR322 and in λ-bacteriophage, screened these with synthetic oligonucleotide probes, subcloned and sequenced the gene, obtained high levels of expression by transformation into a *B. subtilis* strain lacking the genes for both the neutral protease and the alkaline protease (subtilisin), and developed protocols for rapid purification and assay of the enzyme in the broth (P. Thomas, A. J. Russell & J. Brannigan, unpublished results).

This work duplicates much of that reported previously by Wells *et al.* 1983, so I will merely amplify some of our own targets for protein engineering.

Figure 3 shows an *Eco*RI–*Bam*HI fragment containing the gene cloned into pBR322, and figure 4 shows the derived amino-acid sequence, which is identical to that reported by Wells *et al.* (1983). As they observe, the sequence of the mature enzyme is preceded by a typical 'leader sequence', probably responsible for secretion across the membrane, followed by a 'propeptide' of 72 amino acids which contains an unusually high content of charged residues. The Tyr–Ala bond that links this pro-peptide to the enzyme is a typical substrate for subtilisin, so it is probable that this bond is cleaved autocatalytically by previously secreted enzyme.

If the model of subtilisin is rotated so that the active site lies in the NW quadrant, facing the viewer, the charge–relay triad of Asp-32, His-64 and Ser-221 lies just below the S atom of Met-222 at the bottom of a wedge-shaped cleft in the otherwise almost spherical molecules.

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

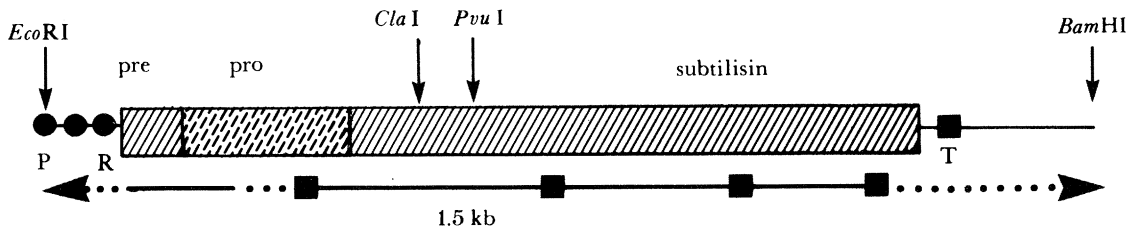

FIGURE 3. The *B. amyloliquefaciens* subtilisin gene. The structure is as proposed by Wells *et al.* (1983). The solid arrows indicate sequence of the gene cloned independently by P. Thomas & J. Brannigan (unpublished data), using the synthetic oligonucleotide fragments (shown as black boxes). P, promoter; R, ribosome binding site; T, terminator.

pre-peptide (32 amino acids)

Met *Arg* Gly *Lys* *Lys* Val Trp Ile Ser Leu Leu Phe Ala Leu Ala Leu

Ile Phe Thr Met Ala Phe Gly Ser Thr Ser Ser Ala Gln Ala Ala Gly–

pro-peptide (72 amino acids)

Lys Ser Asn Gly *Glu* *Lys* *Lys* Tyr Ile Val Gly Phe *Lys* Gln Thr Met

Ser Ala Ala *Lys* *Lys* *Lys* *Asp* Val Ile Ser *Glu* *Lys* Gly Gly *Lys* Val

Gln *Lys* Gln Phe *Lys* Tyr Val *Asp* Ala Ala Ser Ala Thr Leu Asn *Glu*

Lys Ala Val *Lys* Glu Leu *Lys* *Lys* *Asp* Pro Ser Val Ala Tyr Val *Glu*

Glu *Asp* *His* Val Ala *His* Ala Tyr-

B. Amyloliquefaciens subtilisin (275 amino acids)

Ala Gln Ser Val ---------- Ala Ala Gln

FIGURE 4. The amino-acid sequence of *B. amyloliquefaciens*, 'pre-pro-subtilisin' (Wells *et al.* 1983).

The right-hand side of this cleft contains the surface chain of residues 125–127 whose peptide bonds hydrogen bond to the N-terminal side of the incoming peptide substrate, allowing its side chains to protrude into the cleft (Kraut *et al.* 1971). The surface of this cleft is an obvious target for protein engineering to modify substrate specificity on the N-terminal side.

Moreover it is attractive to speculate that the enzyme might initially fold as an inactive zymogen in which at least a portion of the 72-residue 'pro-peptide' would fill the cleft, preventing substrate access. There are many charged residues in this 'pro-peptide', and there are many surface charges in the enzyme that might form buried ion pairs with some of these. If this is so, we must be very careful to avoid substitutions at the surface of the molecule which might impair the folding of the pro-enzyme.

For this reason a model of pro-subtilisin would be most valuable, but there could be great problems in isolating this protein even if the cloned gene is expressed in a host in which the natural protease genes are deleted, as is our strategy. The protein would probably auto-activate. Hence we have constructed a mutant in which Ala replaces the active site Ser-221, which would obviously lack all protease activity. Preliminary experiments show that negligible amounts of this mutant protein are secreted compared with the massive secretion of wild-type protein,

suggesting that auto-activation may be essential for release from the cell membrane or cell wall. If so, we shall be forced to attempt theoretical modelling of the putative zymogen to guide our protein engineering strategies.

Another obvious target for protein engineering of subtilisin is to improve stability by introducing new disulphide bridges. To this end we have been exploring the strategies listed above, by use of molecular graphics in collaboration with the I.B.M. Research Laboratory in Winchester. Only two candidates for isosteric internal substitutions emerged; a pair of hydrogen-bonded residues (Ser-24–Ser-87); and Met-199, whose methyl group is in contact with Gly-178. There are far more candidates for introducing new surface disulphide bridges, but we are wary of these for the reason already discussed.

An even more obvious target is to reduce the susceptibility of the enzyme to inactivation by mild oxidizing agents, including molecular oxygen, by modifying methionine residues. Two of these, Met-50 and Met-222, lie at the surface and it is obvious from both model building and from experiment (Stauffer & Etson 1969) that oxidation of the latter would prevent substrate access. We have made one mutant gene incorporating a Gln codon to replace Met-222, but the far more extensive work reported by Dr Wells (this symposium) makes this study redundant.

This example reinforces my conviction that it is possible to use enzymes of commercial relevance as models for protein engineering and at the same time extend the boundaries of basic science. The properties of the library of Met-222 mutants prepared by the Genentech group have much to tell us about enzyme mechanisms, but they are also likely to produce an enzyme detergent that will work in the presence of bleaching agents!

References

Blow, D. M., Birktoft, J. J. & Hartley, B. S. 1969 *Nature, Lond.* **221**, 337–340.
Eveleigh, D. E. 1981 *Scient. Am.* **245**, 120–133.
Kraut, J., Robertus, J. D., Birktoft, J. J., Alden, R. A., Wilcox, P. E. & Powers, J. C. 1971 *Cold Spring Harb. Symp. quant. Biol.* **36**, 117–123.
Perry, L. J. & Wetzel, R. 1984 *Science, Wash.* **226**, 555–557.
Perutz, M. F. & Raidt, H. 1975 *Nature, Lond.* **255**, 256–259.
Robertus, J. D., Kraut, J., Alden, R. A. & Birktoft, J. J. 1972 *Biochemistry* **11**, 4293–4296.
Stauffer, C. E. & Etson, D. 1969 *J. biol. Chem.* **244**, 5333–5338.
U.S. Department of Commerce 1984 *High technology industries: biotechnology.*
Wells, M. A., Ferrari, E., Henner, D. J., Estell, D. A. & Chen, E. Y. 1983 *Nucl. Acids Res.* **11**, 7911–7924.
Wilkinson, A. J., Fersht, A. R., Blow, D. M., Carter, P. & Winter, G. 1984 *Nature, Lond.* **307**, 187–188.

Discussion

N. L. Brown (*Department of Biochemistry, University of Bristol*). Professor Hartley correctly identifies determination of the crystallographic structure as the slowest stage in applying site-directed mutagenesis to many proteins. However, for many applications of the type that he has described, would not localized but random mutagenesis, such as bisulphite treatment of a single-stranded region of the gene, be sufficient? This would generate mutant proteins, and these could subsequently be screened to identify the ones with the desired properties.

B. S. Hartley. One can certainly generate mutants rapidly by random mutagenesis of a cloned gene, but expression, isolation and screening of the properties of a particular product are the major task. In some cases, clever genetic tricks might be used to distinguish active from inactive

products, but the specific improvements that are commercially relevant will seldom be accessible to this approach.

R. W. Carrell (*University of Otago, New Zealand*). The availability of genetically engineered therapeutic products has highlighted our relative ignorance of the relation of the molecular structure of proteins to their physiology within the body. A specific example is the role of oligosaccharide side chains. For commercial purposes it is much easier to produce glycoproteins without their carbohydrate component. Our experience with α_1-antitrypsin is that the lack of carbohydrate side chains does not affect its function but does markedly decrease its stability and, in particular, its biological half life (Travis *et al.* 1985). Does Professor Hartley agree that if we are to embark on the rational design of proteins for therapeutic use, as well as understanding the relation between molecular structure and function, we also need to know more about the relation between molecular structure and physiology? In particular we need to know more about the role of the oligosaccharide side chains, the factors that determine biological half life, and the contribution of the presence or absence of carbohydrate to protein antigenicity.

Reference

Travis, J., Owen, M. C., George, P., Carrell, R. W., Rosenberg, S., Hallewell, R. A. & Barr, P. J. 1985 *J. biol. Chem.* **260**, 4384–4389.

B. S. Hartley. It is indeed becoming a rule that glycosidic side chains seldom affect the specific function of a protein but are most important in the biological half life, particularly clearance from serum. Hence I strongly support Dr Carrell's argument.

Mary Waye (*M.R.C. Laboratory of Molecular Biology, Cambridge*). What is known about the proteolytic properties of enzymes? One difficulty in over-expressing proteins in *E. coli* is that some foreign proteins are readily degraded in *E. coli*, for example, yeast histone 3 (Mellado & Murray 1983). What approach can we take to engineer proteins that are not susceptible to proteolysis?

Reference

Mellado, R. P. & Murray, K. 1983 *J. molec. Biol.* **168**, 489–503.

B. S. Hartley. We probably knew more about the structure and activity of proteolytic enzymes than any other class of enzyme. Unfortunately the proteases of *E. coli* are an exception to this statement, and more work is justified.

The best approach to avoid proteolysis is to engineer proteins that are very stable to overall (or local) denaturation, because proteolytic enzymes attack only flexible structures.

M. F. Perutz (*M.R.C. Laboratory of Molecular Biology, Cambridge*). Professor Hartley said that the mechanism of denaturation of proteins by extremes of pH was unknown and that it was therefore difficult to engineer resistance to such extremes. In fact the mechanism of denaturation of proteins by alkali emerged from a comparison of the amino-acid sequences of three haemoglobins: adult and foetal human haemoglobin and bovine haemoglobin (Perutz 1974). In 0.1 M NaOH at 20 °C, adult and foetal human carbon monoxyhaemoglobins are denatured

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

with half times of 20 s and 12 min respectively, while bovine haemoglobin is stable. It takes 0.5 M NaOH to denature it, with a half time of 3 h. Human adult haemoglobin contains two buried cysteines and one buried tyrosine, human foetal haemoglobin contains only one buried cysteine and no buried tyrosine, while bovine haemoglobin contains neither buried cysteines nor tyrosines. The sulphydryl groups of cysteine have a microscopic pK of 9.1–9.5 and the phenolic groups of tyrosine have a pK of 9.5–10.0. These groups might have their pKs raised because they lie buried, but they would certainly become ionized above pH 12. Ionization of these groups will attract water of solvation into the non-polar interior of the protein and initiate denaturation. The absence of buried groups that are ionizable at high pH must be responsible for the remarkable resistance to alkali in bovine haemoglobin. Denaturation of proteins by acid is probably activated by ionization of buried histidines. Therefore, to engineer alkali-stable proteins, buried cysteines and tyrosines should be avoided; and for acid-stable proteins, buried histidines should be avoided.

Reference

Perutz, M. F. 1974 *Nature, Lond.* **247**, 341–344.

B. S. Hartley. I am very aware of Dr Perutz's point, because I recall many discussions with him in the early 1970s in which I made this argument. I regret that I did not quote the specific reference that helps to establish it.

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