Subject Name: Microbiology

Semester: IV

Name of the Teacher:

Dr. Kuntal Kanti Goswami

Topic:

Applications of Recombinant DNA Technology

(CC10; Unit 6)

B.Sc (HONOURS) MICROBIOLOGY (CBCS STRUCTURE)

CC-10: Applications of Recombinant DNA Technology (THEORY)

SEMESTER -IV

Products of human therapeutic interest:

Insulin:

Insulin, synthesized by the β-cells of the islets of Langerhans in the pancreas, controls the level of glucose in the blood. An insulin deficiency manifests itself as diabetes mellitus, a complex of symptoms which may lead to death if untreated. Fortunately, many forms of diabetes can be alleviated by a continuing program of insulin injections, thereby supplementing the limited amount of hormone synthesized by the patient's pancreas.

Several strategies have been used to obtain recombinant insulin. One of the first projects, involving synthesis of artificial genes for the A and B chains followed by production of fusion proteins in *E. coli*, illustrates a number of the general techniques used in recombinant protein production.

Synthesis and expression of artificial insulin genes:

In the late 1970s, the idea of making an artificial gene was extremely innovative. Oligonucleotide synthesis was in its infancy at that time, and the available methods for making artificial DNA molecules were much more cumbersome than the present-day automated techniques. Nevertheless, genes coding for the A and B chains of insulin were synthesized as early as 1978.

The procedure used was to synthesize trinucleotides representing all the possible codons and then join these together in the order dictated by the amino acid sequences of the A and B chains. The artificial genes would not necessarily have the same nucleotide sequences as the real gene segments coding for the A and B chains, but they would still specify the correct polypeptides. Two recombinant plasmids were constructed, one carrying the artificial gene for the A chain, and one the gene for the B chain.

In each case the artificial gene was ligated to a lacZ′ reading frame present in a pBR322-type vector. The insulin genes were therefore under the control of the strong lac promoter, and were expressed as fusion proteins, consisting of the first few amino acids of β-galactosidase followed by the A or B polypeptides. Each gene was designed so that its β-galactosidase and insulin segments were separated by a methionine residue, so that the insulin polypeptides could be cleaved from the β-galactosidase segments by treatment with cyanogen bromide. The purified A and B chains were then attached to each other by disulphide bond formation in the test tube.

Synthesis of human growth hormones in *E. coli*:

At about the same time that recombinant insulin was first being made in *E. coli*, other researchers were working on similar projects with the human growth hormones somatostatin and somatotrophin. These two proteins act in conjunction to control growth processes in the human body, their malfunction leading to painful and disabling disorders such as acromegaly (uncontrolled bone growth) and dwarfism.

Somatostatin was the first human protein to be synthesized in *E. coli*. Being a very short protein, only 14 amino acids in length, it was ideally suited for artificial gene synthesis. The strategy used was the same as described for recombinant insulin, involving insertion of the

artificial gene into a lacZ′ vector, synthesis of a fusion protein, and cleavage with cyanogen bromide.

BT transgenic:

Research is being carried out on a wide range of insecticidal proteins from diverse sources. However, all commercially produced insect-resistant transgenic crops express toxin proteins from the Gram-positive bacterium *Bacillus thuringiensis* (Bt). Unlike other *Bacillus* species, Bt produces crystals during sporulation, comprising one or a small number of ~130 kDa protoxins called crystal proteins. These proteins are potent and highly specific insecticides. Approximately 150 distinct Bt toxins have been identified and each shows a unique spectrum of activity. Bt toxins have been used as topical insecticides since the 1930s, but never gained widespread use, because they are rapidly broken down on exposure to daylight and thus have to be applied several times during a growing season. Additionally, only insects infesting the exposed surface of sprayed plants are killed. These problems have been addressed by the expression of crystal proteins in transgenic plants. Bt genes were initially introduced into tomato (Fischhoff et al. 1987) and tobacco (Barton et al. 1987, Vaeck et al. 1987) and later into cotton (Perlak et al. 1990), resulting in the production of insecticidal proteins that protected the plants from insect infestation. However, field tests of these plants revealed that higher levels of the toxin in the plant tissue would be required to obtain commercially useful plants (Delannay et al. 1989). Attempts to increase the expression of the toxin gene in plants by use of different promoters, fusion proteins and leader sequences were not successful. However, examination of the bacterial cry1Ab and cry1Ac genes indicated that they differed significantly from plant genes in a number of ways (Perlak et al. 1991). For example, localized AT-rich regions resembling plant introns, potential plant polyadenylation signal sequences, ATTTA sequences that can destabilize mRNA and rare plant codons were all found. The elimination of

undesirable sequences and modifications to bring codon usage into line with the host species resulted in greatly enhanced expression of the insecticidal toxin and strong insect resistance of the transgenic plants in field tests (Koziel et al. 1993). By carrying out such enhancements, Perlak and colleagues expressed a modified cry3A gene in potato to provide resistance against Colorado beetle (Perlak et al. 1993). In 1995, this crop became the first transgenic insect-resistant crop to reach commercial production, as NewLeafTM potato marketed by Monsanto. The same company also released the first commercial transgenic, insect-resistant varieties of cotton (BollgardTM, expressing cry1Ac and protected against tobacco bollworm) and maize

(YieldGardTM, expressing cry1Ab and resistant to the European corn-borer). Many other biotechnology companies have now produced Bt-transgenic crop plants resistant to a range of insects. Although Bt-transgenic plants currently dominate the market, there are many alternative insecticidal proteins under investigation. Two types of protein are being studied in particular: proteins that inhibit digestive enzymes in the insect gut (proteinase and amylase inhibitors) and lectins (carbohydratebinding proteins). Research into these alternatives is driven in part by the fact that some insects are not affected by any of the known Bt crystal proteins. Homopteran insects, mostly sap-sucking pests such as planthoppers, fall into this category, but have been shown to be susceptible to lectins such as Galanthus nivalis agglutinin (GNA). This lectin has been expressed in many crops, including potato (Shi et al. 1994, Gatehouse et al. 1996), rice (Bano- Maqbool & Christou 1999), tomato and tobacco (reviewed by Schuler et al. 1998).

Antisense Molecules:

Antisense RNA and the engineering of fruit ripening in tomato

Commercially grown tomatoes and other soft fruits are usually picked before they are completely ripe, to allow time for the fruits to be transported to the marketplace before they

begin to spoil. This is essential if the process is to be economically viable, but there is a problem in that most immature fruits do not develop their full flavor if they are removed from the plant before they are fully ripe.

Using antisense RNA to inactivate the polygalacturonase gene

The time scale for development of a fruit is measured as the number of days or weeks after flowering. In tomato, this process takes approximately eight weeks from start to finish, with the color and flavor changes associated with ripening beginning after about six weeks. At this time a number of genes involved in the later stages of ripening are switched on, including one coding for the polygalacturonase enzyme. This enzyme slowly breaks down the polygalacturonic acid component of the cell walls in the fruit pericarp, resulting in a gradual softening. The softening makes the fruit palatable, but if taken too far results in a squashy, spoilt tomato.

Partial inactivation of the polygalacturonase gene should increase the time between flavor development and spoilage of the fruit. To test this hypothesis, a 730 bp restriction fragment was obtained from the 5′ region of the normal polygalacturonase gene, representing just under half of the coding sequence. The orientation of the fragment was reversed, a cauliflower mosaic virus promoter was ligated to the start of the sequence, and a plant polyadenylation signal attached to the end. The construction was then inserted into the Ti plasmid vector pBIN19. Once inside the plant, transcription from the cauliflower mosaic virus promoter should result in synthesis of an antisense RNA complementary to the first half of the polygalacturonase mRNA. This would be sufficient to reduce or even prevent translation of the target mRNA.

Transformation was carried out by introducing the recombinant pBIN19 molecules into *Agrobacterium tumefaciens* bacteria and then allowing the bacteria to infect tomato stem segments. Small amounts of callus material collected from the surfaces of these segments were

tested for their ability to grow on an agar medium containing kanamycin (pBIN19 carries a gene for kanamycin resistance). Resistant transformants were identified and allowed to develop into mature plants. The effect of antisense RNA synthesis on the amount of polygalacturonase mRNA in the cells of ripening fruit was determined by northern hybridization with a single stranded DNA probe specific for the sense mRNA. These experiments showed that ripening fruit from transformed plants contained less polygalacturonase mRNA than the fruits from normal plants. The amounts of polygalacturonase enzyme produced in the ripening fruits of transformed plants were then estimated from the intensities of the relevant bands after separation of fruit proteins by polyacrylamide gel electrophoresis, and by directly measuring the enzyme activities in the fruits. The results showed that less enzyme was synthesized in transformed fruits. Most importantly, the transformed fruits, although undergoing a gradual softening, could be stored for a prolonged period before beginning to spoil. This indicated that the antisense RNA had not completely inactivated the polygalacturonase gene, but had nonetheless produced a sufficient reduction in gene expression to delay the ripening process as desired. The Genetically Modified (GM) tomatoes—marketed under the trade name "FlavrSavr"—were one of the first genetically engineered plants to be approved for sale to the public, first appearing in supermarkets in 1994.

Gene Therapy:

Gene therapy is any procedure used to treat disease by modifying the genetic information in the cells of the patient. There are two basic approaches to gene therapy: germline therapy and somatic cell therapy. In germline therapy, a fertilized egg is provided with a copy of the correct version of the relevant gene and re-implanted into the mother. If successful, the gene is present and expressed in all cells of the resulting individual. Germline therapy is usually carried out by

microinjection of a somatic cell followed by nuclear transfer into an oocyte, and theoretically could be used to treat any inherited disease.

Somatic cell therapy involves manipulation of cells, which either can be removed from the organism, transfected, and then placed back in the body, or transfected *in situ* without removal. The technique has most promise for inherited blood diseases (e.g., hemophilia and thalassaemia), with genes being introduced into stem cells from the bone marrow, which give rise to all the specialized cell types in the blood. The strategy is to prepare a bone extract containing several billion cells, transfect these with a retrovirus-based vector, and then re-implant the cells. Subsequent replication and differentiation of transfectants leads to the added gene being present in all the mature blood cells. The advantage of a retrovirus is that this type of vector has an extremely high transfection frequency, enabling a large proportion of the stem cells in a bone marrow extract to receive the new gene.

Somatic cell therapy also has potential in the treatment of lung diseases such as cystic fibrosis, as DNA cloned in adenovirus vectors or contained in liposomes is taken up by the epithelial cells in the lungs after introduction into the respiratory tract via an inhaler. However, gene expression occurs for only a few weeks, and as yet this has not been developed into an effective means of treating cystic fibrosis. With those genetic diseases where the defect arises because the mutated gene does not code for a functional protein, all that is necessary is to provide the cell with the correct version of the gene: removal of the defective genes is unnecessary. The situation is less easy with dominant genetic diseases, as with these it is the defective gene product itself that is responsible for the disease state, and so the therapy must include not only addition of the correct gene but also removal of the defective version. This requires a gene delivery system that promotes recombination between the chromosomal and vector-borne versions of the gene, so that the defective chromosomal copy is replaced by the gene from the

vector. The technique is complex and unreliable, and broadly applicable procedures have not yet been developed.

Recombinant Vaccine:

The greatest success with recombinant vaccine has been with hepatitis B virus. Hepatitis B is endemic in many tropical parts of the world and leads to liver disease and possibly, after chronic infection, to cancer of the liver. A person who recovers from hepatitis B is immune to future infection because their blood contains antibodies to the hepatitis B surface antigen (HBsAg), which is one of the virus coat proteins. This protein has been synthesized in both Saccharomyces cerevisiae, using a vector based on the 2 µm plasmid, and in Chinese hamster ovary (CHO) cells. In both cases, the protein was obtained in reasonably high quantities, and when injected into test animals providedprotection against hepatitis B.

The key to the success of recombinant HbsAg as a vaccine is provided by an unusual feature of the natural infection process for the virus. The bloodstream of infected individuals contains not only intact hepatitis B virus particles, which are 42 nm in diameter, but also smaller, 22 nm spheres made up entirely of HBsAg protein molecules. Assembly of these 22 nm spheres occurs during HbsAg synthesis in both yeast and hamster cells and it is almost certainly these spheres, rather than single HbsAg molecules, that are the effective component of the recombinant vaccine. The recombinant vaccine therefore mimics part of the natural infection process and stimulates antibody production, but as the spheres are not viable viruses the vaccine does not itself cause the disease. Both the yeast and hamster cell vaccines have been approved for use in humans, and the World Health Organization is promoting their use in national vaccination programmes.

Protein Engineering and Site directed mutagenesis:

One of the most exciting aspects of recombinant DNA technology is that it permits the design, development and isolation of proteins with improved operating characteristics and even completely novel proteins. The simplest example of protein engineering involves site-directed mutagenesis to alter key residues, as originally shown by Winter and colleagues (Winter et al. 1982, Wilkinson et al. 1984). From a detailed knowledge of the enzyme tyrosyltRNA synthetase from *Bacillus stearothermophilus*, including its crystal structure, they were able to predict point mutations in the gene that should increase the enzyme's affinity for the substrate ATP. These changes were introduced and, in one case, a single amino acid change improved the affinity for ATP by a factor of 100. Using a similar approach, the stability of an enzyme can be increased. Thus Perry and Wetzel (1984) were able to increase the thermostability of T4 lysozyme by the introduction of a disulphide bond. However, although new cysteine residues can be introduced at will, they will not necessarily lead to increased thermal stability (Wetzel et al. 1988).

Improving enzymes: subtilisin as a paradigm for protein engineering:

Proof of the power of gene manipulation coupled with the techniques of in vitro (random and sitedirected) mutagenesis as a means of generating improved enzymes is provided by the work done on subtilisin over the past 15 years (for review, see Bryan 2000). Every property of this serine protease has been altered, including its rate of catalysis, substrate specificity, pH-rate profile and stability to oxidative, thermal and alkaline inactivation. In the process, well over 50% of the 275 amino acids of subtilisin have been changed. At some positions in the molecule, the effects of replacing the usual amino acid with all the other 19 natural amino acids have been evaluated. Many of the changes described above were made to improve the ability of subtilisin to hydrolyse protein when incorporated into detergents. However, serine proteases can be used

to synthesize peptides and this approach has a number of advantages over conventional methods (Abrahmsen et al. 1991). A problem with the use of subtilisin for peptide synthesis is that hydrolysis is strongly favoured over aminolysis, unless the reaction is undertaken in organic solvents. Solvents, in turn, reduce the half-life of subtilisin. Using site-directed mutagenesis, a number of variants of subtilisin have been isolated with greatly enhanced solvent stability (Wong et al. 1990, Zhong et al. 1991). Changes introduced included the minimization of surface changes to reduce solvation energy, the enhancement of internal polar and hydrophobic interactions and the introduction of conformational restrictions to reduce the tendency of the protein to denature. Designing these changes requires an extensive knowledge of the enzyme's structure and function. Chen and Arnold (1991, 1993) have provided an alternative solution. They utilized random mutagenesis combined with screening for enhanced proteolysis in the presence of solvent (dimethyl formamide) and substrate (casein).

The engineering of subtilisin has now gone one step further, in that it has been modified such that aminolysis (synthesis) is favoured over hydrolysis, even in aqueous solvents. This was achieved by changing a serine residue in the active site to cysteine (Abrahmsen et al. 1991). The reasons for this enhancement derive mainly from the increased affinity and reactivity of the acyl intermediate for the amino nucleophile. These engineered 'peptide ligases' are in turn being used to synthesize novel glycopeptides. A glycosyl amino acid is used in peptide synthesis to form a glycosyl peptide ester, which will react with another C-protected peptide in the presence of the peptide ligase to form a larger glycosyl peptide.