Course- B.Sc. (Botany Honours), Part -3

Paper-VI (Group-B), Molecular Biology

Topic- Artificial Synthesis of Genes.

PDF

Binod Kumar Pandey

Assistant Professor & HOD, Botany Department,

Ram Ratan Singh College, Mokama, Patna.

E.mail- binodkp14@gmail.com

Artificial Synthesis of Genes

For the first time in 1955, Michelson chemically synthesized a dinucleotide in laboratory. Later on in 1970, Har Govind Khorana and K.L. Agarwal for the first time chemically synthesized gene coding for tyrosine tRNA of yeast. For the synthesis of tRNA and rRNA there are specific genes.

However, genes of tRNA are the smallest genes containing about 80 nucleotides. In 1965, Robert W. Holley and coworkers worked out first the molecular structure of yeast alanine tRNA. This structure lent support to Khorana in deduction of structure of the gene.

A gene is responsible for encoding mRNA and mRNA for polypeptide chain. If the structure of a polypeptide chain is known, the structure of mRNA from genetic code dictionary and in turn the structure of gene can easily be worked out. There are two approaches for artificial synthesis of the gene, by using chemicals and through mRNAs.

I. Synthesis of a Gene for Yeast alanine tRNA:

The molecular structure of yeast alanine tRNA was worked out by R.W. Holley and coworkers in 1965, this information helped Khorana to deduce the structure of alanine tRNA. They worked out that yeast alanine tRNA contains 77 base pairs. It was very difficult to assemble 77 base pairs of nucleotides in ordered form.

Therefore, they synthesized chemically the short deoxynucleotide sequences which was joined by hydrogen bonding to form a long complementary strand. By using polynucleotide ligase the double stranded pieces were produced.

The complete procedure of synthesizing gene for yeast alanine tRNA is discussed in the following steps:

1. Synthesis of Oligonucleotides:

In the first approach, fifteen oligonucleotides ranging from penta-nucleotide (i.e. oligodeoxynucleotide of five bases) to an icosanucleotide (i.e. oligodeoxynucleotide of twenty bases) were synthesized.

The chemical synthesis was brought about through condensation between the -OH group at 3′ position of one deoxynucleotide and the $–PO_4$ group at 5′ position of the second deoxynucleotide. All other functional groups of deoxyribonucleotides not taking part in condensation processes were protected so that the condensation could be brought about.

The deoxynucleotides were protected as below:

(i) The amino group of deoxyadenosine was protected by benzyl group (BZ), the amino group deoxyguanosine was protected by isobutyl group. These protective groups were removed by treating with ammonia when synthesis was over.

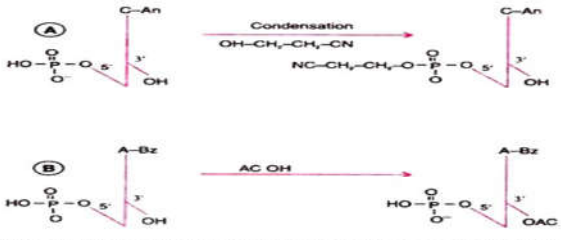

Fig. 6.6 : Protection of 5′ –OH group due to condensation of cyanoethyl group (A) and 3′–OH group due to condensation of acetyl group (B) alanine–tRNA.

(ii) The hydroxyl group (-OH) at the 5′ position of receiving deoxynucleotide was protected by cyano-ethyl group ($HN\text{-}CH_2\text{-}CH_2$-) (Fig. 6.6A).

(iii) The -OH group at the 3′ position of second incoming deoxynucleotide was protected by using acetyl group (Ac) (Fig. 6.6B).

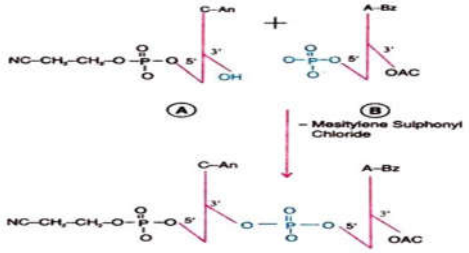

Fig. 6.7 : Condensation of two nuclectides and production of an oligonucleotide.

The different protecting groups used and treatment required to remove the protecting groups are given in Table 6.2. When the groups of deoxynucleotides were protected, the products of Fig. 6.7A and B reacted to form deoxy-oligonucleotide as given in Fig. 6.7.

When deoxynucleotides were condensed into oligonucleotides, different protecting groups were removed by treating with ammonia, acid or alkali (Table 6.2). For example, both the cyanoethyl group at the 5′ position and the acetyl group at the 3′ position were removed by alkali treatment.

Finally, condensation between the groups of two, three or four nucleotides was brought about. The receiving segment had a free 3′-OH group and a protected 5′-OH group, whereas the incoming segment had a free 5′-OH group and a protected 3′-OH group. After each addition, the protective group at the 3′ end had removed so that frees 3′-0H group could receive another segment.

2. Synthesis of Three Duplex Fragments of a Gene:

By using 15 single stranded oligonucleotides, three large double stranded DNA fragments were synthesized (Fig. 6.7).

These three fragments contained:

(i) Segment of A having the first 20 nucleotides with the nucleotides 17-20 as the single stranded,

(ii) Segment B consisting of nucleotides 17-50 with the nucleotides 17-20 and 46-50 as the single stranded, and

(iii) Segment C containing the nucleotides 46-77 with the single stranded region 46-50.

3. Synthesis of a Gene from Three Duplex Fragments of DNA:

The three segments (A, B, C) synthesized as above were joined by using the enzyme polynucleotide ligase to produce the complete gene for alanine tRNA (Fig. 6.8).

The joining of the three fragments was done by any of two following methods:

(i) In one approach, fragment A was joined to B by taking advantage of overlapping in nucleotide residues 17-20. Then, the fragment C was added with the overlap in nucleotides 46-50. Thus, a complete double stranded DNA with 77 base pairs was prepared.

(ii) In the second approach, the fragment B was joined to C. At the end the fragment A was added to nucleotide residues 17-20 to obtain the complete gene for alanine tRNA. Khorana (1970) prepared this gene in vitro which was used for future work.

They found that alanine tRNA gene replicated and transcribed into tRNA just like the natural gene. It is not known whether tRNA prepared from artificially synthesized gene had the molecular organisation similar to alanine tRNA or not.

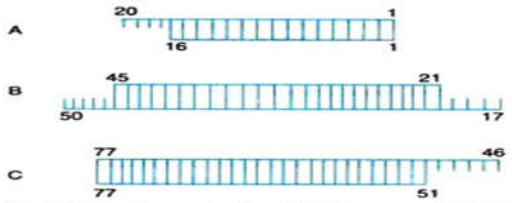

Fig. 6.8: The three duplex DNA fragments (A, B and C) were synthesised for the synthesis of the gene for yeast alanine tRNA.

II. Artificial Synthesis of a Gene for Bacterial Tyrosine tRNA:

In 1975, Khorana and co-workers completed the synthesis of a gene for E. coli tyrosine tRNA precursor. E. coli tRNA precursors are formed from the larger precursors. The tyrosine tRNA precursor has 126 nucleotides. They synthesized the complete sequence of DNA duplex coding for tyrosine-tRNA precursor of E. coli. Though these segments are not the proper structural gene yet are the regions involved in its regulation.

Twenty six small oligonucleotide DNA segment giving rise to tRNA precursor was synthesized which were arranged into six double stranded fragments each containing single stranded ends. These six fragments were joined to give rise to complete gene of 126 base pairs for tyrosine tRNA precursor of E. coli.

Khorana (1979) completely synthesized a biologically functional tyrosine tRNA suppressor gene of E. coli which was 207 base pairs long and contained (i) a 51 base pairs long DNA corresponding to promotor region, (ii) a 126 base pair long DNA corresponding to precursor region of tRNA, (iii) a 25 base pair long DNA including 16 base pairs contained restriction site for Eco RI.

This complete synthetic gene was joined in phage lambda vector which in turn was allowed to transfect E. coli cells. After transfection phage containing synthetic gene successfully multiplied in E. coli. Khorana (1979) made the phosphodiester approach for synthesizing the oligonucleotides of the biologically active tRNA.

The demerits of this approach are:

(i) The completion of reaction in long time,
(ii) Rapidly decrease in yield with the increase in chain length, and
(iii) Time taking procedure of purification.

III. Artificial Synthesis of a Human Leukocyte Interferon Gene:

Interferons are proteinaceous in nature produced in human to inhibit viral infection.

These are of three types secreted by three genes i.e:

(i) Leukocyte interferon gene (IFN-α gene),
(ii) Fibroblast interferon gene (IFN-β gene) and
(iii) Immune interferon gene (IFN-γ gene).

In 1980, Weisman and coworkers published the nucleotide sequence of IFN-α gene. Taking advantage of this information, Edge (1981) successfully synthesized the total human interferon gene of 514 base pairs long.

Edge (1981) made the phosphotriester approach in artificially synthesizing 67 oligonucleotides of 10-20 nucleotide residue long segments. The phosphotriester approach over-comes some of the demerits of phosphodiester approach by blocking the function of each intemucleotide phosphodiester during the process of synthesis. A completely protected mononu-cleotide containing a fully masked 3′ phosphate triester group is used in this method.

Coupling of initial nucleotide onto a polyacrylamide resin was done to which further nucleotides in pairs were added. In this way 66 oligonucleotides of 14-21 nucleotide residues

were first synthesized. These were arranged in predetermined ways and joined chemically. The 514 base pair long IFN-α gene contained the initiation and termination signals.

Edge (1981) incorporated the artificially synthesized gene into a plasmid through biotechnological technique. The recombinant plasmid was transferred into E.coli cells which expressed a-interferon. This technique now-a-days is being adopted to produce interferon commercially.

IV. Gene Machine:

Recently, fully automated commercial instrument called automated polynucleotide synthesizer or gene machine is available in market which synthesizes predetermined polynucleotide sequence.

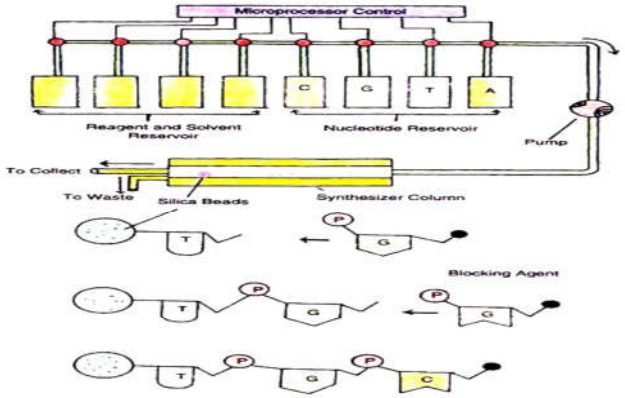

Fig. 6.9 : A gene machine and its working mechanism (diagrammatic)

Therefore, the genes can be synthesized rapidly and in high amount. For example, a gene for tRNA can be synthesized within a few days through gene machine. It automatically synthesizes the short segments of single stranded DNA under the control of microprocessor.

The working principle of a gene machine includes:

(i) Development of insoluble silica-based support in the form of beads which provides support for solid phase synthesis of DNA chain, and

(ii) Development of stable deoxyribonucleoside phosphoramidites as synthons which are stable to oxidation and hydrolysis, and ideal for DNA synthesis.

The mechanism of a gene machine is shown in Fig. 6.9. Four separate reservoirs containing nucleotides (A, T, C and G) are connected with a tube to a cylinder (synthesiser column) packed with small silica beads.

These beads provide support for assembly of DNA molecules. Reservoirs for reagent and solvent are also attached. The whole procedure of adding or removing the chemicals from the reagent reservoir in time is controlled by microcomputer control system i.e. microprocessor.

If one desires to synthesise a short polynucleotide with a sequence of nucleotides T, G, C, the cylinder is first filled with beads with a single ‘T’ attached. Thereafter, it is flooded with ‘G’ from the reservoir. The right hand side of each G is blocked by using chemicals from the reservoir so that its attachment with any other Gs can be prevented.

The remaining Gs which could not join with Ts are flushed from the cylinder. The other chemicals are passed from the reagent and solvent reservoirs so that these can remove the blocks from G which is attached with the T. In the same way this cycle is repeated by flooding with C from reservoir into the cylinder. Finally, the sequence TGC is synthesized on the silica beads which are removed chemically later on.

The desired sequence is entered on a key board and the microprocessor automatically opens the valve of nucleotide reservoir, and chemical and solvent reservoir. In the gene machine

the nucleotides are added into a polynucleotide chain at the rate of two nucleotides per hour. By feeding the instructions of human insulin gene in gene machine insulin has been synthesized.

V. Gene Synthesis by Using mRNA:

The mRNAs are the transcripts of genes that have to be translated into polypeptides or proteins. It is rather difficult to identify a particular gene on a chromosome. However, it is easier to pick up the mRNAs and synthesize a gene because the total population of mRNA in a cell remains high.

The majority of mRNA contains a long poly adenylated tract i.e. about 100 (A) residues at 3′ terminus. Therefore, the mRNA can be separated from the rest of RNA population of the cell.

The mRNAs are passed through an oligo-dT cellulose affinity column (Fig. 6.10). The poly (A) binds to (T). The oligo-dT segment contains 10-20 nucleotides which hybridizes the poly (A) of mRNA. The oligo-dT provides primer at poly (A) region with a free 3′-OH end. The reverse transcriptase uses the free end and synthesizes a single stranded cDNA in the presence of dCTP, dGTP, dATP and dTTP, and results in mRNA-cDNA hybrid.

At the end the enzyme forms a loop by using the last few bases as the template. This results in synthesis of a short 'hairpin' loop at 3'end of the cDNA. The mRNA is degraded from mRNA-cDNA hybrid by using alkali. This phenomenon is known as hydrolysis. Consequently mRNA is separated.

In the next step the single stranded DNA acts as template for the synthesis of double stranded DNA in the presence of polymerase I and all the four deoxynucleotides. The hairpin acts as primer for chain elongation. Finally, a double stranded cDNA is synthesized. SI nuclease is used to cleave hairpin loop and result in double stranded cDNA (Fig. 6.10).

In 1970, S. Mizutami, H.M. Temin and D. Baltimore discovered the RNA dependent DNA polymerase i.e. reverse transcriptase in retroviruses. This enzyme yields single stranded DNA on RNA template. For the discovery of reverse transcriptase R. Dulbecco, Temin and Baltimore were awarded Nobel Prize in 1975.

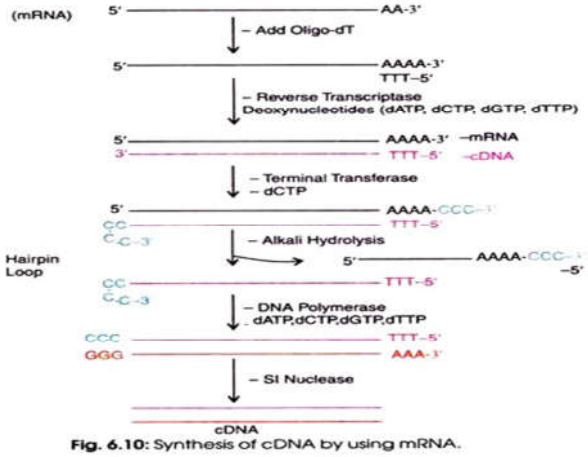

Fig. 6.10: Synthesis of cDNA by using mRNA.

Land (1981) have given an improved method for cDNA synthesis. The single stranded cDNA complementary to mRNA is tailed with oligo-dC tail. This process is facilitated by using the enzyme terminal transferase and adding dCTP nucleotide. The tailing is followed by oligo-dG priming of the second strand synthesis. This checks the formation of hairpin loop. Therefore, SI nuclease is not required.