Nucleotides recombination states

J.B.Savkovic-Stevanovic

Abstract: - DNA sequencing strands information were defined. The synthesis rates of purine and pyrimidine oxyribonucleotides and deoxyribonucleotides characterized DNA features. The most important regulator of de novo purine biosynthesis is the intracellular concentration of the phosphoribosyl-pyrophosphate-phosphoribosylamine. Mammalian liver is a major site of purine nucleotide synthesis and provides purines in the form of bases or nucleosides to be salvaged and utilized by those tissues incapable of synthesizing purines de novo. In this paper DNA recombination was studied. The information value of the fragment in genome was derived.

Key-Words: - Recombination**,** enzyme, nucleotide, DNA features, genomic library, bioinformation.

I. INTRODUCTION

The structure and function of the purines and pyrimidines and their nucleosides and nucleotides were studied in numerous literature. Synthetic analogs of naturally occurring nucleotides find application in cancer chemotherapy as enzyme inhibitors and can replace the naturally occurring nucleotides in nucleic acids. Therapeutic attempts to inhibit the growth of cancer cells or certain viruses have often employed administration of analogs of bases, nucleosides, or nucleotides that inhibit the synthesis of either DNA or RNA. Allopurinol, a purine analog, is widely used in the treatment of gout.

Biomedical important it neither nucleotides nor their parent purine and pyrimidine bases in the diet are incorporated into human tissue nucleic acids or into purine or pyrimidine coenzymes. Even when a diet rich in nucleoproteins is ingested, human subjects form the constituents of tissue nucleic acids from amphibolic intermediates. This de novo synthesis permits purine and pyrimidine analogs with potential as anticancer drugs to be incorporated into DNA. The rates of synthesis of purine and pyrimidine oxy- and deoxyribonucleotides are subject to precise regulation[1]-[3]. Mechanisms have evolved to ensure production of these compounds in quantities and at times appropriate to meet varying physiologic demand[4]-[7].

Recombinant DNA technology, often referred to as genetic engineering, has revolutionized biology and is having an ever-increasing impact on clinical medicine. Much has been learned about human genetic disease from pedigree analysis and study of affected proteins, but in many cases where the specific genetic defect is unknown, these approaches can not be used. The approach circumvents these limitations by going directly to the DNA molecule for information.

Understanding recombinant DNA approach is important for several reasons. The information explosing occuring in this area is truly staggering. To understand and keep up with this field, one must have an appreciation of the fundamental concepts involved. There is now rational approach to understanding the molecular bases of a number of diseasese.g.familial hypercholesterolemia, sickle cell disease, the thalassemias, cystic fibrosis, Huntington's chorea. Using recominant DNA technology, human proteins can be produced in abundance for therapy(e.g. insulin, growth hormone, plasminogen activator). Proteins for vaccines, e.g. hepatitis B, and for diagnostic tests ,e.g. AIDS test, can be obtained. Recominant DNA technology is used to diagnose existing diseases and predict the risk of developing a given disease. Gene therapy for sickle cell disese, the thalassemias, adenosine deaminase deficiency, and other diseases may be devised. Gene therapy has already been accomplished in mice, in whom hereditary hypogonadism has been corrected by transgenic injection of the gonadotropin releasing hormone gene into the fertilized ovum.

In this paper recombinant mechanisms of the nucleotides synthesis were studied.

II. SYNTHETIC NUCLEOTIDES

Synthetic analogs of nucleobases, nucleosides, and nucleotides are widely used in the medical sciences and clinical medicine. In the past, most of these uses have depended upon the role of nucleotides as components of nucleic acids for cellular growth and division. For a cell to divide, its nucleic acids must be replicated. This requires that the precursors of nucleic acids – the normal purine and pyrimidine deoxy- be readily available. One of the most important components of the oncologict's pharmacopeia is the group of synthetic analogs of purine and pyrimidine nucleobases and nucleosides.

The pharmacologic approach has been to use an analog in which either the heterocyclic ring structure of the sugar moiety has been altered in such a way as to induce toxic effects when the analog becomes incorporated into various cellular constituents. Many of these effects results from inhibition by the drug of specific enzyme activities necessary for nucleic acids synthesis or from the incorporation of metabolites of the drug into the nucleic acids where they alter the base pairing essential to accurate transfer of information.

Examples of these would be the 5-fluoro or 5-iodo derivatives of uracil or deoxyuridine, which serve as thymine or thymidine analogs, respectively. Borh 6-thioguanine and 6-mercaptopurine, in which naturally occuring hydroxyl groups are replaced with thiol groups at the 6 position, are widely used clinically. The analogs in which the purine or pyrimidine ring contains extra nitrogen atoms, such as 5- or 6- azauridine or azacytidine and 8-azaguanine, also have been tested clinically.

The purine analog 4-hydroxypyrazolopyrimidine (allopurinol) is widely marketed as an inhibitor of de novo purine biosynthesis and of xanthine oxidase. It is used for the treatment of hyperuricemia and gout. Nucleosides containing arabinose rather than ribose as the sugar moieties, e.g. cytarabine (arabynosil cytosine-Ara-C), are used in chemotherapy of cancer and viral infestions.

Azathiopirine, which is catabolized to 6-mercptopurine, is useful in organ transplantion as a supressor of events involved in

immunologic rejection. A series of nucleoside analogs with antivirial activities has been studied for several years, one 5-iododeoxyuridine, is effective in the local treatment of herpetic keratitis, an infection of the cornea by herpesvirus.

Numerous analogs of purine and pyrimidine ribonucleotides have been synthesized so as to generate nonhydrolyzable di- or triphosphates for use in vitro. These analogs allow the investigator to determine wheither given biochemical effects of nucleoside di-or triphosphates require hydrolysis or wheither their effects are mediated by occupying specific nucleotide binding sites on enzymes or regulatory proteins.

The reduction of ribonucleoside diphosphates to deoxyribo-nucleoide diphosphates is subject to complex regulation.

III.CATABOLISM OF PURINE

In humans the ultimate catabolite (end product) of purines is uric acid. On the basis of observations in humans with inherited enzyme deficiencies, it appears that over 99% of the uric acid is derived from substrates of purine nucleoside phosphorylase, a component of the purine salvage pathway. The purine products of purine nucleoside phosphorilase, guanine and hypoxanthine, are converted to uric acid by way of xanthine in reaction catalysed by the enzymes guanase and xanthine oxidase, respectively, of liver, small intestine, and kidney.

Xanthine oxidase is an important site for pharmacologic intervention in patients with hyperuricemia and gout. In lower primates and other mammals, but not in humans, uricase hydrolyses uric acid to allantoin, a highly water soluble end product. Ambhibians, birds, and reptiles, which lack uricase, excrete uric acid and guanine as the end products of both purine metabolism and nitrogen (protein) metabolism.

IV. CATABOLISM OF PYRIMIDINES

The catabolism of pyrimidines, which occurs mainly in the liver, produces highly soluble end products. This contrasts with the production of the sparingly soluble uric acid and sodium urate by purine catabolism. The release of respiratory CO_2 from the ureido carbon of the pyrimidine nucleus represents a major pathway for the catabolism of uracil, cytosine, and thymine. β-alanine and β-aminoisobutyrate are the major end products of cytosine , uracil, and thymine catabolism, respectively.

Thymine is the precursor β-aminoisobutyrate in laboratory animals and in humans. The excretion of β-aminoisobutyrate is increased in leukemia as well as after the body has been subjected to x-irradiation. This is undoubtedly a reflection of increased destruction of cells and their DNA. A familial occurnce of an abnormally high excretion of β-aminoisobutyrate has also been observed n otherwise normal individuals. This genetic trait is traceable to recessively expressed gene. Higs excretors results only when the trait is homozygous. Approximately 25% of tested . Although little is known about the mechanism whereby β-aminoisobutyrate is degraded in humans, an enzyme that catalyzes the reverible transamination reaction has been identified in pig kidney. The β-aminoisobutyrate is convereted to methylmalonic semialdehyde and thence to propionate, which in turn proceeds to succinate.

The initial steps in degradation of pyrimidine nucleotides, including the removal of the sugar phosphate moitely by hydrolysis of the N-glucosidic bond are similar to reversing the later part of the synthetic pathway. For pseudouridine, which is formed in t-RNA by a rearrangement reaction, there is no mechanism to catalyze the hydrolysis or phosphorylsis of this unusual nucleoside to its respective pyrimidine base, uracil. Consequently, pseudouridine is excreted unchanged in the urine of normal persons. Some individuals with urate overexctretion greater than 600mg/uric acid per 24 hours can be categorized as having secondary hyperuricemia. They have other diease processes such as cancer or psoriasis that lead enhanced tissue turnover.

Finally, there are persons with identifiable enzyme defects, including abnormalities of PRPP synthetase, the HGPRT ase (hypoxanthine-guanine phosphoribosyl- transferase) deficiencies, both the complete Lesch-Nyhan syndrome and incomplete deficiencies and glucose-6-phosphatase deficiency von Gierke's disease. There exists also a group of patients exhibiting idiopathic overproduction hyperuricemia, which will certainly be regarded as a heterogeneous group of disease ons the molecular bases for their metabolic defects are recognized.

V. DNA PAIRING

DNA is a complex biopolymer that is organized as a double helix[8]. The fundamental organizational element is the sequence of purine (adenine [a} or guanine [G] and pyrimidine (cytosine [C] or thymine [T]) bases (Fig.1). These bases are attached to the C-1' position of the sugar deoxyribose, and the bases are linked together through joining of the sugar moieties at their 3'and 5'positions via a phosphordiester bond. The alternating deoxyribose phosphate groups form the backbone of the double helix. These 3'-5' linkages also define the orientation of a given strand of the DNA molecule, and since the 2 strands run in opposite directions, they are said to be antiparralel.

Basic pairing is one of thr most fundamental concepts of DNA structure and function. Adenine and thymine always pair, by hydrogen bonding, as do guanine and cytosine. These base pairs are said to be complementary, and the guanine content of a fragment of double stranded DNA will always equal its cytosine content, likewise the thymine and adenine contents are equal. Interaction can be reduced by heating the DNA to denaturate it. Base pairing and hydrophobic base stacking interactions hold the 2 DNS strand together. The laws of base pairing predict that 2 complementary DNA strands will reanneal exactly in register upon renaturaton,as happen when the temperature of the solution is slowly reduced to normal. Indeed, the degree of base-pair matching or mismatching can be estimated from the temperature required for denaturation – renaturation. Segments of DNA with high degrees of base – pair matching require more energy input heat to accomplished denaturation, or, to put it another way, a closely matched segment will withstand more heat before the strand separate. This reaction is used to determine whether there are significant differences between 2 DNA sequences, and it underlines the concept of hybridization, which is fundamental to the processes which be described hear.

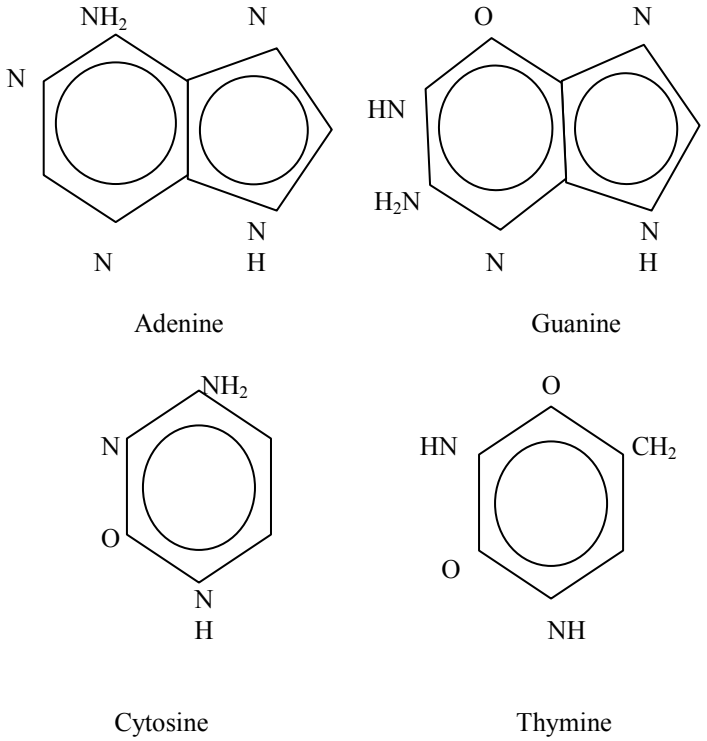

Fig.1 The major four the purine and pyrimidine bases

There are about 3 x 10^9 base pairs (bp) in each human haploid genome. If an average gene length is 3 x 10^3 bp(three kilobases, kb), the genome could consist of 10^6 genes, assuming that there is no overlap and that transcription proceeds in only one direction.

The double – helical DNA is packaged into a more compact structure by a number of proteins, most notably the basic proteins called *histones*. The DNA presents within nucleus of a cell, if simply extended, would be about a meter long. The chromosomal proteins compact this long length of DNA so that it can be packaged into a nucleus with a volume of a few cubic microns.

Gene organization is very interesting. In general, prokaryotic genes consist of a small regulatory region (100-500 bp) and a large protein-coding segment (500-10000 bp). Several genes are often controlled by a single regulatory unit. Most mammalian genes are more complicated, in that the coding region are interrupted by noncoding regions that that are eliminated when the primary RNA transcript is processed into mature messenger RNA (mRNA). The coding regions (those regions that appear in the mature RNA species) are called *exons*, and noncoding regions, which interpose or intervene between the exons, are called *introns*. Introns always removed from precursors RNA before transport into the cytoplasm occurs. The process by which introns are removed from precursor RNA and by which exons are ligated together is called RNA splicing. Incorrect processing of the primary transcript into the mature mRNA can results in disease in humans. Regulatory regions for specific eukaryotic genes are usually located in the DNA that flanks the transcription initiation site at its 5'end (5' flanking –sequence DNA). Occasionally, such sequences are found within the gene itself or in the region that flanks the 3' end of the gene. In mammalian cells, each gene has its own regulatory region. Many eukaryotic genes, and some viruses that replicate in mammalian cells, have special regions, called enhancers, that increase the rate of transcription. Some genes also have DNA sequences, known as silencers, that diminish transcription. Mammalian genes are obviously complicated, multicomponent structure.

In gene transcription, information generally flows from DNA to mRNA to protein as illustrated in Fig.2. This is rigidly controlled process involving a number of complex steps, each of which no doubt is regulated by one or more enzymes or factors, faulty function at any of these steps can cause disease.

Concept use in recombinant DNA technology consists from isolation and manipulation of DNA, including end-to-end joining of sequences from very different sources to make chimeric molecules(eg. molecules containing both human and bacterial DNA sequences in a sequence independent fashion.

VI. CONCEPT USE IN RECOMBINANT DNA

Concept use in recombinant DNA technology consists from isolation and manipulation of DNA, including end-to-end joining of sequences from very different sources to make chimeric molecules(eg. molecules containing both human and bacterial DNA sequencies in a sequence independent fashion. This involves several unique techniques and reagents.

VII RESTRICTION ENZYMES

Certain endonucleses, enzymes that cut DNA as specific DNA sequences within the molecule, as opposed to exonucleases, which digest from the ends of DNA molecules, are a key tool in recombinant DNA research. These enzymes were originally called restriction enzymes because their presence in a given bacterium restricted the growth of certain bacterial viruses called bacteriophages. Restriction enzymes cut DNA into short pieces in a sequence specific manner, in contrast to most other enzymatic, chemical, or physical methods, which break DNA randomly. These defensive enzymes, over 200 have been discovered, protect the host bactyerial DNA from DNA from foreign organisms, primarly infective phages. However, they are only present in cells that also have a companion enzyme that methylates the host DNA, rendering it an unsuitable substrate for digestion by the restriction enzyme.Thus, site-specific DNA methylases and restriction enzymes always exist in pairs in a bacterium.

Restriction enzymes are named after the bacterium from which they are isolated are EcoRI from *Escherichia coli*, Bam HI (Table 1) from *Bacilius amyloliquefaciens.* If the nucleotides are distributed randomly in a given DNA molecule, one can calculate how frequently a given enzyme would cut a length of DNA. For each position in the DNA molecule there are 4 possibilities A,C,G,T, therefore, a restriction enzyme that recognizes a 4 bp sequence will cut, on average, once every 256 bp (4^4), whereas another enzyme that recognizes a 6 –bp sequence will cut once every 4^6 (4096)bp)A given piece of DNA will have a characteristic linear array of sites for the various enzymes, hence, a restriction map can be constructed.

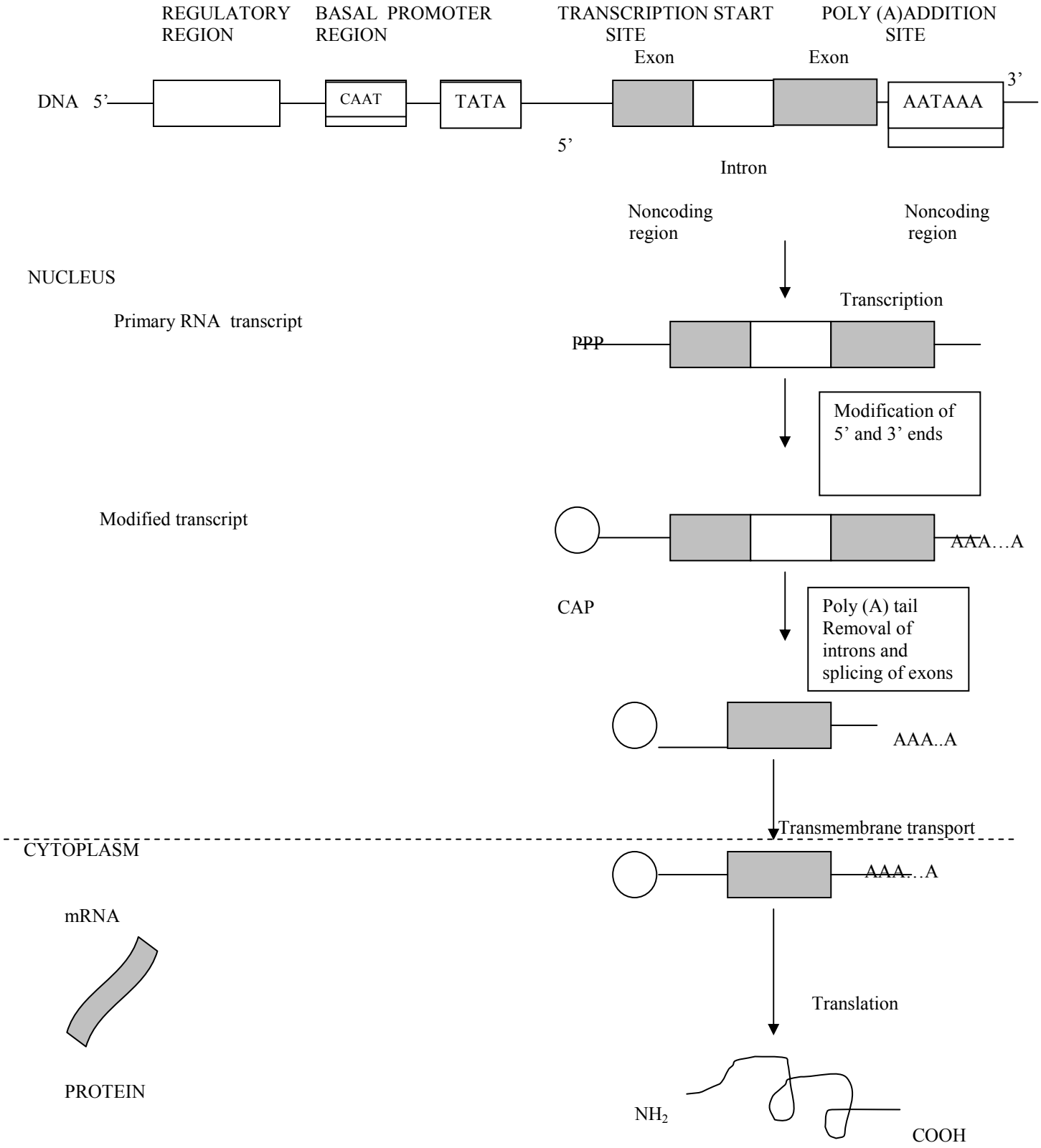

Fig.2 Organization of a eukaryotic transcription and the pathway of eucaryotic gene expression
Eucaryotic genes have structural and regulatory regions. Eucaryotic genes have structural and regulatory regions.The structural region consists of the coding DNA and 5' and 3' noncoding DNA sequences. The coding regions are divided into 2 parts: (1) exons, which eventually become mature RNA, and (2) introns, which are processed out of the primary transcript. The structural region is bounded at its 5' end by the transcription initiation ite and at its 3' end by the polyadenylate addition or termination site. The promotor region , which contains specific DNA sequences that interact with various protein factors.

Tabela 1. Selected restriction endonucleses and their sequence specificities

Endonuclese	Sequence cleaved	Bacterial sources
Bam HI	G G A T C C C C T A G G	*Bacillus amyloliquefaciens H*
Bgl II	A G A T C T T C T A G A	*Bacillius globigii*
Eco RI	G A A T T C C T T A A G	*Eschericia coli RY13*
Eco RII	C C T G G G G A C C	*Eschericia coli R245*
Hind III	A A G C T T T T C G A A	*Haemophilus influenzae* R_d
Hha I	G C G C C G C G	*Haemophilus haemolyticus*
Hpa I	G T T A A C C A A T T G	*Haemophilus parainfluenzae*
Mst II	C C T N A G G G G A N T C C	*Microcoleus strain*
Pst I	C TA G C A G G A C G T C	*Providencia stuartii 164*
Taq I	T C G A A G C T	*Thermus aquaticus YTI*

A-adenine, C- cytosine, G- guanine, T- tymine. Arrows show the site of cleavage, depending on the site, sticky ends (Bam HI) or blunt ends (Hpa I) may result. the lenght of the recognition sequence can be 4 bp (Taq I), 5bp (Eco RII), 6 bp (Eco RI) or 7bp (Mst II). By convention, these are written in the 5' to 3' direction for the upper strand of each recognition , and the lower strand is shown with the opposite (i.e 3' to 5') polarity.

Note that most recognition sequences are palindromes (i.e. the sequence reads the same in oppsite directions on the two strands). A residue designated N means that any nucleotide is permitted.

When DNA is digested with a given enzyme, the ends of all the fragments will have the same DNA sequence.

The primary transcript has a special structure, a cap, at the 5' end and astretch of A's at the 3' end.

The fragments produced can be isolated by electrophoresis on agarose or polyacrylamide, this is an essential step in cloning and a major use of these enzymes.

A number of other enzymes that act on DNA and RNA are an important part recombinant DNA technology. Many of these are reffered to in Table 2.

Sticky end ligation is technically easy, but some special techniques are often required to overcome problems inherent in this approach. Sticky ends of a vector may reconnect with themselves, with no net gain of DNA. Sticky ends of fragments can also anneal, so that tandem heterogenous inserts form. Also, sticky-end sites may not be avilable or in a covnenient position. to circumvent these problems, an enzyme that generates blunt ends is used, and new ends are added using the enzyme terminal transferase. If poly d(G) is added to the 3' ends of the vector and polyd© is dded to the 3' ends of the foreign DNA, the 2 molecules can only anneal to each other, thus circumventing the problems listed above. This procedure called homopolymer tailing, also generates an SmaI restriction site, and so it is easy to retrieve the fragment. Sometimes, synthetic oligonucleotide linkers with a convenient restriction enzyme sequence are ligated to the blunt ended DNA. Direct blunt end ligation is accomplished using the enzyme bacteriophage T4 DNA ligase. This technique, though more difficult than sticky endligation, has the advantage of joining together any pairs of end. The advantages are that there is no control over the orientation of insertion or the number of molecules annealed together, and there is no easy way of retrieving the insert.

VII. CLONING

A clone is a large population of identical molecules, bacteria or cells that arise from a common ancestor. Cloning allows for the production of a large number of identical DNA molecules, which can then be characterized or use for other purposes. This technique is based on the fact that chimeric or hybrid DNA molecules can be constructed in cloning vector, typically bacterial plasmids, phages, or cosmids, which continue to replicate in a host cell under their own control systems. In this way, the chimeric DNA is amplified.

Bacterial plasmid are small, circular duplex DNA molecules whose natural function is to confer antipiotic resistance to the host cell. Plasmids have several properties that make them extremely useful as cloning vectors. They exist as a single or multiple copies within the bacterium and replicate independently from the bacterial DNA.The complete DNA sequence of many plasmids is known, hence, the precise location of restriction enzyme cleavage sites for inserting the for DNA is available. Plasmids are smaller than the host chromosome and are therefore easily separated from the latter, and the desired DNA is readily removed by cutting the plasmid with the enzymes specific for the restriction site into which the original place of DNA ws inserted.

Phages usually have linear DNA molecules into which foreign DNA can be inserted at several restriction enzyme sites. The chimeric DNA is collected after the phage proceeds through its lytic cycle and produces mature, infective phage particles. A major advantage of phage vectors is that while plasmids accept DNA pieces about 6-10kb long, phages can accept DNA fragments 10-20 kb long, a limitation imposed by the amount of DNA that can be packed into the phage head.

Even larger fragments of DNA can be cloned in cosmids, which combine the best features of plasmids and phages. Cosmids are plasmids that contain the dNA sequences, so-called cos sites, required for packing lambda DNA into the phage particle. These vectors grow in the plasmid form in bacteria, but since much of the unnecessary lambda DNA has been removed, more chimeric can be packaged into the particle head. It is not unusual for cosmids to carry inserts of chimeric DNA that are 35-50 kb long. A comparison of these vectors is shown in Table 3.

Insertion of DNA into a functional region of the vector will interfere with the action of thia region, and so care must be taken not to interrupt an essential function of the vector. This concept can be exploited, however, to provide a selection technique. the common plasmid vector pBR322 has both trtra cycline (tet) and ampicillin (amp) resistance genes. a single Pst I site within the amp-reistance gene is commonly used as the insertion site for a piece of foreign DNA. In addition to having sticky ends (Table 1 and Fig.3), the DNA inserted at this site disrupts the ampresistance gene and makes the bacterium carrying this plasmid amp-sensitive (Fig.4). Thus, the parental plasmid, which provides resistance to both antibiotics, can be readily separated from thechimeric plasmid, which is resistant only to tetracycline. Additional confirmation that insertion has taken place comes from sizing the plasmid DNA obtained ffrom the putative recombinant on an agarose gel, since the chimeric DNA molecule is now larger than host vector DNA.

VIII. GENOMIC LIBRARY CONSTRUCTION

The combination of restriction enzymes and various cloning vectors allows the entire genome of an organism to be packed into a vector. A collection of these different recombinant clones is called library. A genomic library is prepared from the total DNA of a cell line or tissue. A cDNA library represents the population of mRNAs in a tissue. Genomic libraries are prepared by performing partial digestion of total DNA with a restriction enzyme that cuts DNA frequently. Phage vectors are preferred for these libraries because they accept large pieces of DNA up to 20 kb. The goal is to achieve a complete library. The number of fragments required to attain this objective is inversely related to fragment size and directly related to genome size.

The The composition of some complete genomic libraries are shown in the Table 4.

Tabela 2. Enzymes used in recombinat DNA

Enzyme	Reaction	Primary use
Alkaline phosphatase	Dephosphorylates 5' ends of RNA and DNA	Removal of 5' $-PO_4$ groups prior to kinase labeling to self-ligation.
Bal 31 nuclease	Degrades both the 3' and 5' ends of DNA	Progressive shortening of DNA molecules
DNA ligase	Catalyses bonds between DNA molecules	Joining of DNA molecules
DNA polymerase I	Synthesizes double stranded DNA from single stranded DNA	Synthesis of double stranded cDNA, nick translation.
Dnase I	Under appropriate conditions, produces single stranded nicks in DNA	Nick translation, mapping of hypersensitive sites.
Econuclease III	Removes nucleotides from 3' ends of DNA	DNA sequencing, mapping of DNA –protein interactions.
λ Econuclease	Removes nucleotides from 5' ends of DNA	DNA sequencing
Polynucletide kinase	Transfer terminal phosphate (γ position) from ATP to 5' –OH groups of DNA or RNA	^{32}P labeling of DNA or RNA
Reverse transcriptase	Synthesi DNA from RNA templatezes	Synthesis of cDNA from mRNA, RNA(5' ends) –mapping studies.
Sl nuclease	Degrades single stranded DNA	Removal of "hairpin" in synthesis of cDNA, RNA r studies (both 5' and 3' ends).
Terminal transferase	Adds nucleotides to the 3' ends of DNA	Homopolymer tailing

Table 3. Common cloning vectors

Vector	DNA insert size
Plasmid pBR322	0.01-10kb
Lambda charon 4A	10-20kb
Cosmids	35-50kb

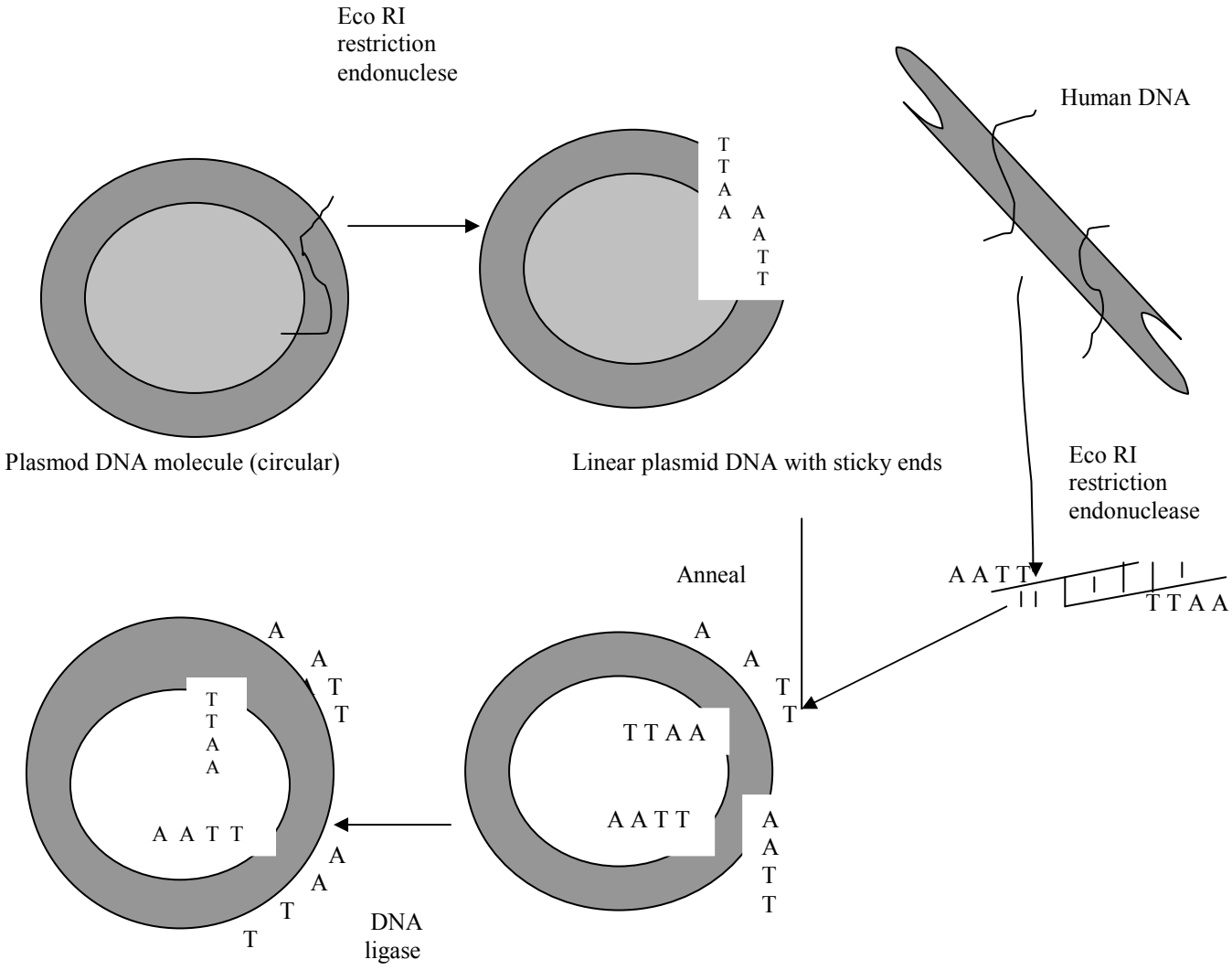

Fig.3 Use of restriction nucleases to make new recombinant or chimeric DNA molecules

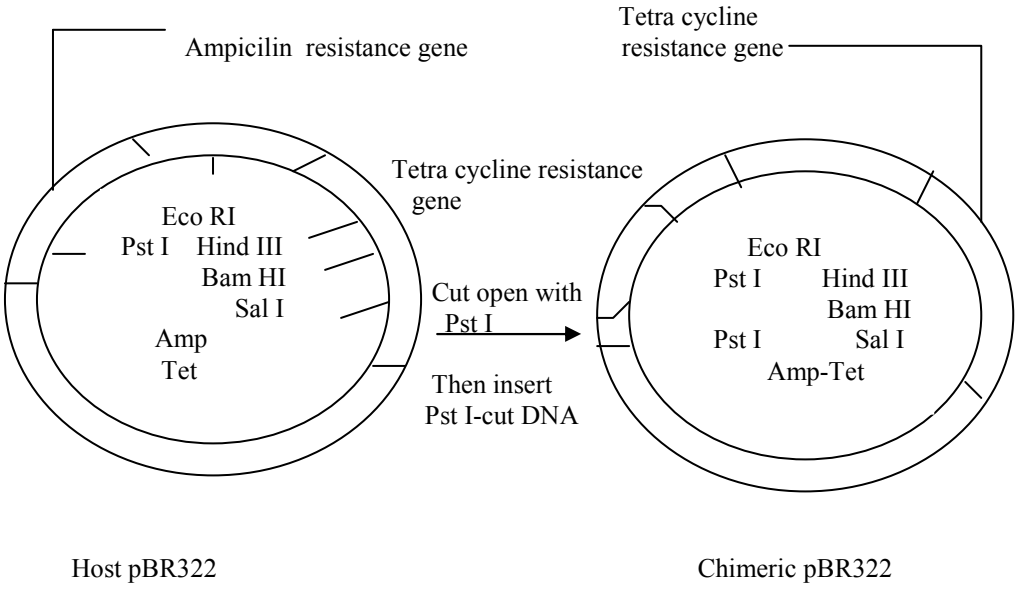

Fig.4 A method of screening recombinants for inserted DNA fragments

Table 4. The composition of the complete genomic libraries

Source	Complete genomic library
E coli	1500 fragments
Yeast	4500 fragments
Drosophila	50000 fragments
Mammals	800000 fragments

The number of random fragments, unique clones that a library should have to ensure that any single gene is represented is inversely related to the average fragment size used to construct the library and directly related to the number of genes in the organism. The numbers given above represent the number of fragments, independent clones, necessary to achieve 99% probability of finding of a given DNA sequence in a recombinant DNA library with an average insert of 5x 10^3 nucleotides. The differences represent the variations in genomic complexity between the creatures.
The number of clones necessary N,

$$N = \frac{\ln(1-P)}{\ln(1-f)} \qquad (1)$$

where P is the probability desired and f is the function of the total genome in a single clone.

In the case of mammalian genomic library given the presence of $3x10^9$ nucleotides in the haploid genome, the equation is becoming:

$$N = \frac{\ln(1-0.99)}{\ln[1-(\frac{5x10^3}{3x10^9})]} = 1$$

Information value will be defined [9]:

$$I_v = -\sum_i^N P_i \log P_i \qquad (2)$$

where P_i probability of finding *i- th* DNA sequence.

cDNA libraries are prepared by first isolating all the mRNAs in a tissue and then copying these molecules into double stranded DNA, using, sequentially, the enzymes reverse transcriptase and DNA polymerase. Plasmids are often the favored vectors for cDNA libraries because they are much more convenient to work with than are phages or cosmids, although
various lambda phage vectors have special advantages for cDNA cloning.

A vector in which the protein coded by the gene introduced by recombinant DNA technology is actually synthesized is known as an expression vector. Such vectors are now commonly used to detect specific cDNA molecules in libraries and to produce proteins by genetic engineering techniques

specific cDNA molecules in libraries and to produce proteins by genetic engineering techniques.

A variety of molecules can be used to "probe" libraries in search f a specific gene of cDNA molecule or to define and quantitate DNA or RNA separated by electrophoresis through various gels. Probes are generally pieces of DNA or RNA labeled with a ^{32}P-containing nucleotide. The probe must recognize a complementary sequence to be effective. A cDNA synthesized from a specific mRNA can be used to screen either a cDNA library for a longer cDNA or a genomic library for a complementary sequence in the coding region of a gene. A popular technique for finding specific gene entails taking a short amino acid sequence and, using the codon usage for that species, making an oligonucleotide probe that will detect the corresponding DNA fragment in a genomic library. If the sequence match exactly, probes 15-20 nucleotides long will hybridize.

Visualization of a specific cDNA or RNA fragment among the many thousands of "contaminating "molecules requires the convergence of a number of techniques, which are collectively termed blot transfer.

Colony or plaque hybridization is the method by which specific clones are identified and purified. Bacteria are grown on colonies on an agar plate and overlade with a nitrocellulose filter paper. Cells from each colony stick to the filter and are permanently fixed thereto by heat, which with sodium hydroxide treatment also lyses the cells and denatures the DNA so that it will hybridize with the probe.

IX. DNA SEQUENCING

The segments of specific DNA molecules obtained by recombinant DNA technology can be analyzed for their nucleotide sequence. This method depends upon having a large number of identical DNA molecules. This requirement can be satisfied by cloning the fragment of interest, using the known techniques.

The isolation of a specific gene from an entire genome requires a technique that will discriminate one part in a million. The identification of a regulatory region that may be only 10bp in length requires a sensitivity of one part in 3 x 10^8, a disease such as sickle cell anemia is caused by a single base change, or one part in 3 x 10^9. Recombinant DNA technology is powerful enough to accomplish all of these things.

Gene mapping is technique which is used to localize specific genes to distinct chromosomes and, to define a map of the human genome. This is already yielding useful information in the definition of human disease. Somatic cells hybridization and in

situ hybridization are 2 techniques used to accomplish this. In in site hybridization, the simpler and more direct procedure, a radioactive probe is added to a metaphase spread of chromosomes on a glass slide. The exact area of hybridization is localized by layering photographic emulsion over the slide and, after exposure lining up the grains with some histologic identification of the chromosome. This often places the gene at a location on given band or region on the chromosome. The following conclusions can already ber drawn: (1) Genes that code for proteins with similar function can be located on separate chromosomes (α and β globin). (2) Genes that form part of a family can also be on separate chromosomes, growth hormone and prolactin. (3) the genes involed in many hereditary disorders known to be due to specific protein deficiencies, including X chromosome linked conditions, are ideed located at pecific sites. Of most interst , perhaps is the fact that because of the availability of defined and cloned restriction fragments, the chromosomal location for many disorders for which the protein deficiency is unknown is being defined

Once the defect is localized to a region of DNA that has the characteristic structure of a gene (Fig.2), a synthetic gene can be constructed and expressed in an appropriate vector and its function can beassessed, or the putative peptide, deduced from the open reading frame in the coding region, can be synthesized. Antobodies directed against this peptide is expressed in normal persons and wether it is absent in those with the genetic syndrome.

There is a normal variation of DNA sequence just as there is with more obvious aspects of human structure. Variations of DNA sequencie, polymorphisms occur approximately once in every 500 nucleotides, or about 10^7times per genome. There are, no doubt deletions and insertions of DNA as well as single base substitutions. In healthy people, these alterations obviously occur in noncoding regions of DNA or at sites that cause no change in function of the encoded protein. This polymorphism of DNA structure can be used to search for the specific gene involved.

X. CONCLUSIONS

Variations of DNA sequence occur approximately about 10^7 times per genome. The isolation of a specific gene from an entire genome need to develope a technique that will discriminate one part in a million. The nucleotides are distributed randomly in a given DNA molecule. In this paper DNA sequence information value was defined.

Carbamyl phosphate synthase is sensitive to feedback inhibition by both purine and pyramidine nucleotides and activation by PRPP.

A practical goal of recombinant DNA research is the production of materials for biomedical application

Notation
f-function of the total genome in a single clone
N-number of clones
P- probability of the sequence appear

Abervation
AMP- adenine monophosphate
ADP- adenine diphosphate
ATP- adenine triphosphate
CTP- cytosine triphosphate
DNA-deoxyribonucleic acid
DP-diphosphate
GDP- guanine diphosphate
GMP-guanine monophosphate
GTP- guanine triphosphate
HGPRTase-hypoxanthine-guanine
phosphoribosyl- transferase
IMP- monophosphate
MP-monnophosphate
NADPH-cofactor
OMP- orotidine monophosphate
PRPP- phosphoribosyl-pyrophosphate
RNA-ribonucleic acid
TMP-thymidine monophosphate
TP- triphosphat
UMP- uridine monophosphate
UTP- uridine triphosphate

References
[1] J.Savkovic-Stevanovic, Human tissue information processing, *WSEAS, 2nd International Conference Biomedical Electronics and Biomedical Informatics* ID 618-272,Moscow, Russia, August 20-22,2009, pp.64-69
[2]J.Savkovic-Stevanovic, Informational macromolecule in biological systems, *MCBC2009-Mathematics and Computers in Biology and Chemistry,* Prague**,** March,23-25,2009.
[3]J.F.Henderson, A.R.P. Paterson, *Nucleotide metabolism:An introduction,* Academic Press,1973.
[4] A.M.Michelson, *The chemistry of nucleosides and nucleotides,* Academic Press,1963. W.H., Prusoff, D.C.Ward, Nucleoside analogs with antivirial activity, *Biochem. Pharmacol. 25,1976,1233.*
[5]B.N.Ames, et al., Uric acid provides an antioxidant defense in humans against oxidant- and radical –caused aging and cancer: A hypothesis. *Proc. Natl. Acad. Sci* USA,78,1981,6858.
[6]J.Savkovic-Stevanovic., Genetic information in biopolymeric chain, World Scientists Engineering and Academic Society-*WSEAS, 2nd International Conference Biomedical Electronics and Biomedical Informatics,* ID618-337,Moscow, Russia, August 20-22, 2009 ,pp. 122-126 .
[7] D.W.Martin,Jr.,Gelfand E.W. Biochemistry of diseases of immunodevelopment, *Annu Rev. Biochem.* 50,1981,845.
[8] B.Lewin, Genes II, Wiley & Sons, N.J.,1985.
[9] J.Savkovic-Stevanovic, *Informatics,* Faculty of Technology and Metallurgy, Belgrade University, Chapter2, 2007.