

Genetic Engineering and Biotechnology

V. Kumar Gera

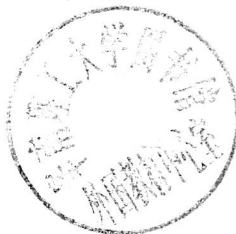


Q78
G354

Genetic Engineering and Biotechnology

V. KUMAR GERA

HoD, Biotechnology
IILM, G. Noida



E2007000081



**DHRUV
PUBLICATIONS**

103, Aggarwal Tower, 1st Floor, Plot No. 2
Pkt. O&P, Dilshad Garden, Delhi-110 095

Cell. : +919811546614

GENETIC ENGINEERING AND BIOTECHNOLOGY

First Edition : 2006

Copyright © Author & Publisher

ISBN. 81-8240-009-0

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means including photocopy, E-mail, Websites, recording or any other storage and retrieval system without the prior written permission from the publisher.

Price : Rs. 650

Published by	: Thomas John for Dhruv Publications
Laser typeset	: Chandra Computers, R-26, Rita Block, Shakarpur, Delhi-92. Ph. : 20098438

Preface

Biotechnology, a contraction of biological technology, can be defined as application of biological organism, system or process to manufacturing and service industries. However, many scientists consider only the techniques of recombinant DNA (r-DNA) that find applications as biotechnology. Genetic Engineering is an umbrella term for a range of techniques for isolation, analyzing and manipulating DNA in vitro that result in predetermined alteration in genotype of an organism. The genetic engineers have been endowed with immense power of producing genetically modified organisms. The r-DNA technology has given a boost to production of novel variety of crop plants, medical research, environmental protection and developments in almost every facet of life that one can think of. The literature has become so vast and deep that no one can follow all the developments in the field. It is in this scenario that technology of genetic engineering as applied to biotechnological processes has been attempted in this text book.

The book deals with the principles of genetic engineering, the strategic developments in the discipline approaches adopted for modification of genetic information content of the organisms, applicability of control systems in metabolic processes and the future prospects of application of these technologies in microbes, plants and animals for benefit of humankind. In such a text repetition of the technology is imperative. However attempt has been made to avoid such repetitions as far as possible but wherever necessary it has been retained. Further an attempt has also been made to incorporate in brief the details of experimental processes to explain the strategies developed and solutions sought.

The book is expected to serve undergraduate and postgraduate students of Biotechnology (B.Sc. and M.Sc.), Engineering, Agricultural and Medical Sciences of various Indian universities. A humble attempt has been made to present the subject in its totality, in a comprehensive lucid manner and all humanly possible care has been taken to avoid mistakes. Still any mistake found in the book is my sole responsibility and for which suggestions of the reader for improvement of the text are most welcome.

I with great indebtedness acknowledge the contributions of scores of scientists who have been dedicated to Biotechnology. I am also indebted to my colleagues and students at Meerut for their unflagging support, intellectual stimulation that they have provided all these many years. I am also thankful to my family members and publishers for their constant support during the preparation of the book.

V. Kumar Gera

G. Noida
September, 2005

CONTENT

GENETIC ENGINEERING AND BIOTECHNOLOGY

1. DNA : The Basic Molecule of 'Gene'

1.1.	Introduction	1
1.2.	Structure of Nucleic Acids	3
1.3.	Base Composition of DNA : Chargaff's Rule	4
1.3.1.	Structure of DNA	4
1.4.	Base Pairing and DNA Architecture	8
1.5.	Configurations of DNA	9
1.5.1.	A-DNA	11
1.5.2.	Z-DNA	11
1.6.	Base-pair Stacking and Flexibility	12
1.7.	Major and Minor Grooves	13
1.8.	Synthesis and Replication of Nucleic Acids	14
1.8.1.	Nucleic Acid Synthesis	15
1.8.2.	DNA Replication	17
1.9.	Initiation of Replication	19
1.10.	Elongation Phase of Replication	21
1.11.	Termination of Replication	22
1.12.	Replication Accuracy	23

2. Gene Cloning and Construction of Genomic Libraries

2.1.	Introduction	25
2.2.	Genetic Engineering	26
2.3.	Basic Concepts in Gene Cloning	27
2.3.1.	Cutting the DNA	28
2.3.2.	Restriction Endonucleases	28
2.3.2.1.	Restriction Modification Phenomenon	29
2.3.2.2.	Types of Endonucleases	30
2.3.2.3.	Restriction Sites	31
2.3.2.4.	Characters and Applications	34
2.3.2.5.	Choice of Restriction Enzymes	37
2.3.3.	Size Selection	38
2.4.	Joining DNA Molecules	39
2.5.	Cloning Vehicles (Vectors)—their nature and characters	45
2.5.1.	Phage Vectors	52
2.5.2.	Cosmids	56
2.5.3.	Single Stranded DNA Phages as Cloning Vectors (M13, fd, and f1 Phages)	57
2.6.	Return of Recombinant DNA into Host Cell	59
2.6.1.	Transformation-Transfection	60
2.6.2.	Transduction	61
2.7.	Screening	62

2.7.1.	Direct Selection	62
2.7.2.	Indirect Selection	63
2.7.2.1.	Nucleic Acid Hybridization	63
2.7.2.2.	Hybrid Arrest Translation (HART) and Hybrid Release Translation (HTR) Selection	66

3. Gene Cloning and cDNA Libraries

3.1.	Introduction	69
3.2.	cDNA and cDNA Bank	70
3.3.	Usefulness of cDNA Clones	71
3.3.1.	Easy Handling and Simple Screening	71
3.3.2.	Accuracy in Screening	72
3.3.3.	When RNA only is Genetic Material	72
3.3.4.	For Special Aims	72
3.4.	Preparation of cDNA	74
3.4.1.	Isolation and Purification of mRNA	74
3.4.	Preparation of cDNA	74
3.4.1.1.	Quantitative Requirements of mRNA for cDNA Synthesis	75
3.4.1.2.	Extraction of Total RNA	76
3.4.1.3.	Extraction of mRNA	78
3.4.2.	cDNA Synthesis	80
3.4.2.1.	Reverse Transcription	80
3.4.2.2.	Formation of Double Stranded cDNA	82
3.4.2.3.	Removal of Hairpin Loop and ss cDNA (SI Nuclease Digestion)	82
3.4.2.4.	Size Fractionation	83
3.4.2.5.	Enrichment of ds-cDNA	83
3.5.	Formation of Cohesive Ends on ds-cDNA	84
3.6.	Transformation of E. coli cells	86
3.7.	Screening of cDNA Clone Bank	87
3.7.1.	Primary Screening of cDNA Clones	87
3.7.2.	Probe	87
3.7.3.	Secondary Screening	88

4. Expression of Foreign DNA in Bacterial Cell

4.1.	Introduction	92
4.2.	Gene Expression	93
4.2.1.	Gene Expression in prokaryotes	94
4.2.2.	Mechanism of Gene Expression in Eukaryotes	96
4.2.3.1.	Ribosome Binding Site/Sequence	98
4.2.3.2.	Promoter Sites and their Recognition	98
4.2.3.3.	Introns and their Processing	99
4.3.	Strategies for Expression of Foreign DNA in Prokaryotes	101
4.3.1.	Fusion Genes	101
4.3.2.	Native Proteins	102
4.3.3.	Case Study of Human Growth Hormone Gene Expression in E. coli (Goeddel et al., 1979)	103

4.4.	Detection of Eukaryotic Gene Expression	106
4.4.1.	Complementation	106
4.4.4.	Gel Electrophoresis	107
4.4.5.	Detection through Specialized Expression Cell System	107
4.4.5.1.	Minicells	107
4.4.5.2.	Maxicells	110
4.5.	Impediments and Solutions to Maximising Foreign DNA Coded Products	111
4.5.1.	Promoter Strength	112
4.5.2.	Usage of Codons	112
4.5.3.	Ribosome Binding Site (AGGA) and the Initiation Codon (AUG)	112
4.5.4.	Transcription Terminator Sequence	113
4.5.5.	Plasmid Copy Number and its Stability	113
4.5.6.	Improving Screening Procedures	113
4.5.7.	Alternative Expression Host System	114

5. Eukaryotic Gene Cloning Systems: I Yeast

5.1.	Introduction	115
5.2.	Technological Strategies for Yeast Cloning	116
5.3.	Transformation Procedure for Yeast.....	116
5.3.1.	Transformation	117
5.3.2.	Selection of the Transformants	118
5.3.3.	Genetic Markers in the Vectors for Selection	118
5.3.3.1.	Modes of Transformation and Types of Yeast Cloning Vectors	119
5.3.3.2.	Integrative Vectors	120
5.3.3.3.	Cut and Gapped Vectors	122
5.3.3.4.	Autonomously Replicating Vectors	123
5.3.3.5.	Other Strategies	128
5.3.3.6.	Expression of Cloned Gene	131
5.3.3.7.	Expression Sequences/Promoters	134
5.3.3.8.	Expression Via Fusion Protein by fusion Plasmid	135
5.3.3.9.	Heterologous Gene Expression	136
5.3.3.10.	Heterologous Proteins and their Secretion/Processing	137
5.4.	Yeast Cloning and Biotechnology	138

6. Eukaryotic Gene Cloning Systems: II. Mammalian Cells

6.1.	Introduction	140
6.2.	Cloning in Mammalian Cells	141
6.2.1.	METHODOLOGY FOR TRANSFECTION	141
6.2.1.1.	Facilitation Agents	141
6.2.2.	Alternate Methods of Transfection	143
6.2.2.1.	Electroporation	144
6.2.2.2.	Protoplast Fusion	145
6.2.2.3.	Micro-injection	145

6.2.2.4.	Biolistic Method	146
6.2.3.	Fate of the Taken-up (transfecting) DNA	147
6.2.4.	Vectors for Cloning in Mammalian Cells	148
6.2.4.1.	Plasmid Based Vectors and Cloning in Mammalian cells	149
6.2.4.2.	Viral Based Vectors	152
6.2.4.3.	Important Viral Vectors	154
6.2.5.	Selection of transformed Cells	160
6.2.5.1.	Thymidine Kinase as Selectable Marker	161
6.2.5.2.	Gpt System as the Selectable Marker	161
6.2.5.3.	Aminoglycoside Transferase System as Selectable Marker	163
6.2.5.4.	Dihydrofolate Reductase as Selectable Marker	164

7. Genetic Engineering and Medicine: I. Prophylaxis

7.1.	Introduction	166
7.2.	Vaccines	166
7.3.	Present Status and Problems in Vaccine(s) Development ...	167
7.3.1.	Inability to Obtain Pathogens in Culture	168
7.3.3.	Problems Related with Inactivated (Killed) Viruses	169
7.3.4.	Problems Due to Variations in Antigenic Determinants	170
7.4.	The Novel Methods for Vaccine Production	171
7.4.1.	Gene Cloning and Expression for Immunogenic Proteins	171
7.4.1.1.	Limitation	171
7.4.1.2.	Other Achievements	172
7.5.	Production of Live Virus Vaccines through Genetic Engineering	172
7.5.1.	Modification of Immunogenic Systems	172
7.5.2.	Genetically Engineered Attenuation	173
7.5.3.	Synthetic Peptides as Vaccines	173
7.5.3.1.	Selection of Peptide Sequence	174
7.5.4.	Recombinant—DNA Mediated Vaccines	176
7.5.4.1.	General Biology of Vaccinia Virus	177
7.5.4.2.	Strategies and Methodology of Production of Recombinant Vaccines	177
7.5.4.3.	Some Examples of recombinant Vaccinia	179
7.5.4.4.	Prospects	183
7.5.4.5.	Problems	185

8. Genetic Engineering and Disease Diagnosis

8.1.	Introduction	186
8.2.	Diagnosis of Genetic Disorder	186
8.2.2.	Thalassaemias	189
8.2.2.1.	Gross Deletions	190
8.2.2.2.	Defective Synthesis	191
8.2.2.3.	Promoter Mutations	191
8.2.2.4.	Defect in post-transcriptional Processing of mRNA	191
8.2.2.5.	Mutations that Alter mRNA Function	192

8.2.3.	Technologies Involved in Diagnosis of Haemoglobinopathies	192
8.2.3.1.	Restriction Enzyme Analysis	193
8.2.3.2.	Use of Synthetic Oligonucleotides	193
8.2.3.3.	Restriction Enzyme Polymorphism	194
8.2.4.	Diagnosis of Other Inheritable Diseases	195
8.2.5.	Diagnostic Analysis for Forensic Use	199
8.2.6.	Diagnosis of Viral Diseases	199

9. Genetic Engineering and Medicines: III. Therapeutics

9.1.	Introduction	201
9.2.	Fermentation Process and Chemotherapeutic Agents:	
	Antibiotics as Therapeutic Agents	201
9.2.1.	Enhancing Productivity	203
9.2.1.1.	Increased Gene Dose of Biosynthetic Genes	203
9.2.1.2.	Improved Promoter Activity	203
9.2.1.3.	Precursor-Product Relationship	203
9.2.1.4.	Elimination of Feedback Control	204
9.2.1.5.	Strain Improvement and Novel Secondary Metabolite Production	204
9.2.2.	Limitation	205
9.3.	Peptides as Therapeutic Agents	206
9.3.1.	General Strategies for Peptide Synthesis	206
9.3.2.	Proteins for Therapeutic Uses	207
9.3.2.1.	Hormones	208
9.3.2.2.	Some Examples of other Hormones	
	Attractive for Genetic Engineers	210
9.3.3.	Supplementation Proteins for Short or Long Term Therapy	211
9.3.3.1.	Tissue Plasminogen Activator	211
9.3.3.2.	Interferon	211
9.3.3.3.	Erythropoietin	213
9.3.4.	Proteins Required for Regulatory Network	214
9.3.4.1.	Immune System and Proteins	214
9.3.4.2.	Proteins for Cell Growth and its Control	216
9.4.	Other Proteins of Interest	217
9.4.1.	Genetically Engineered Non-autologous Proteins	217
9.4.2.	Second Generation Proteins	217
9.5.	Futuristic Approach	218
9.6.	Gene Therapy	219
9.7.	Genomics, Proteomics and Biotechnology	221

10. Genetic Engineering in Plants and Agriculture: I. Microbes and Plants

10.1.	Introduction	224
10.2.	Genetic Engineering and Disease Diagnosis in Plants	226
10.2.1.	Barley Yellow Dwarf Viral Disease (BYDV)	226
10.2.2.	Genetic Probes	226

10.2.3.	Other Diagnostic Probes	227
10.3.	Plant Production and Microbial Associations	228
10.3.1.	Symbiotic and Associative Micro-organisms	228
10.3.1.1.	Biological Nitrogen Fixation	228
10.3.1.2.	Rhizobium-legume Symbiosis	229
10.3.2.	Nitrogen Fixing Bacteria and Exogenous Nitrogen	246
10.3.3.	Transfer of Nitrogen Fixing Genes to Crop Plants	246
10.3.4.	Symbiotic Association of Frankia	247

11. Genetic Engineering in Plants and Agriculture:

II. Gene Technology and Crop Improvement

11.1.	Introduction	248
11.2.	Gene Identification and Isolation	249
11.2.1.	Methods of Gene Isolation	249
11.2.2.	Methods of Gene Identification	251
11.2.2.1.	By Use of Selective Probes	251
11.2.2.2.	By Heterologous Hybridization	252
11.2.2.3.	By Nucleic Acid Hybridization	252
11.2.2.4.	Gene Identification through RFLPs	253
11.2.2.5.	Gene Identification by Mutagenesis through Transposons	253
11.2.2.6.	Gene Identification through Expression	254

12. Genetic Engineering in Plants and Agriculture: III. Transforming Vectors in Plants

12.1.	Introduction	255
12.2.	Crown Gall Bacterium (Agrobacterium) and Ti-plasmid	256
12.2.1.	Molecular Biology of Ti-Plasmids	259
12.2.1.1.	DNA Homology	259
12.2.1.2.	Genetic Map of Ti-plasmid	261
12.2.1.3.	T-DNA: Ti-Plasmid and Development of Gene Vector	262
12.3.	Synthesis/Construction of Ti-plasmid based Gene Vector	267
12.3.2.	Recombinant Plasmid Transfer to A. Tumefaciens	270
12.3.3.	Agrobacterium Infection and Gene Delivery	271
	Case Study(II): Transfer of Phaseolin Gene into Sunflower Crown Gall Cells	273
12.4.	Alternative Gene Vector for Plants (CaMV)	275
12.4.1.	Structure and Function of CaMV Genome	276
12.5.	Methods of Direct Gene Transfer/Gene Delivery	280
12.5.1.	Chemical Treatments to Induce DNA Uptake by Isolated Cells and Protoplasts	280
12.5.2.	Liposome Mediated Uptake	281
12.5.3.	Fusions with the Bacterial Sphaeroplasts	281
12.5.4.	Microinjections of DNA into Plant Nuclei	281
12.5.5.	Macroinjection of DNA to soak Meristems	281
12.5.6.	DNA Transfer by 'Shot Gun' Method	281
12.5.7.	Electroporation	282

13. Genetic Engineering In Plants and Agriculture: IV. Some Cherished goals and Prospects in Plant Improvement through Gene Technology	
13.1. Introduction	284
13.2. Productivity	284
13.3. Nitrogen Metabolism	286
13.4. Control of Plant Development	287
13.4.1. Phytochrome	287
13.4.2. Phytohormones	288
13.4.4. Seed Development and Storage Product: The Yield	289
13.5. Improvement in Nutritional Quality of Food Grains	289
13.5.1. Insertion of Extra-Codons in Genes	290
13.5.2. Modifying Expression of Existing Genes	290
13.5.3. Mutagenesis and Selection	290
13.6. Biochemical Quality and Industrial Requirement	291
13.7. Resistance to Herbicides	292
13.8. Resistance to Insects	294
13.8.1. Bt-Toxins	294
13.8.2. Pheromones	296
13.8.3. Antifeedant Chemicals	297
13.8.4. Host Recognition Compounds	297
13.9. Resistance to Viral Infection	297
13.9.1. Cross-Protection	297
13.9.2. Use of Satellite RNAs	298
13.9.3. Antisense RNA and Protection against Viral Attack	298
14. Genetic Engineering and Microbial Products	
14.1. Introduction	300
14.2. Fermentation and Biotechnology	301
14.2.1. Biomass Production	303
14.2.2. Fermentation Process and Production of Microbial Enzymes	304
14.2.2.1. Biotechnological Genetic Engineering in Beverage Production	304
14.2.2.2. Biotechnological Genetic Engineering in Dairy Products	305
14.2.2.3. Other Enzymes	306
14.3. Biotransformations	312
14.3.1. Techniques of Immobilization of Cell and Enzymes	313
14.3.2. Bioreactors	315
14.4. Genetic Engineering and Microbial Secondary Metabolites	318
14.4.1. Manipulating Genes for Regulation	318
14.4.2. Manipulating Genes for Quality	318
14.4.3. Bioconversion Process	320
14.4.4. Mutational Biosynthesis	320
14.4.5. Specific Enzymic Activities	321

15. Techniques and Technologies

15.1.	Chemical Synthesis of DNA.....	322
15.1.1.	Phosphoramidite Method	323
15.1.1.1.	Uses of Synthesized Oligonucleotides	325
15.1.2.	Synthesis of Genes	326
15.2.	Genomic Sequencing	328
15.3.	Chromosome Walking	330
15.4.	Sequencing of Nucleic Acids	334
15.4.1.	Enzymatic Method for Sequencing of DNA (Sanger's Technique):	335
15.4.1.1.	Principle	335
15.4.1.2.	Technology	337
15.4.2.	Approaches to Sequence Larger Stretches of DNA	340
15.4.3.	Chemical Method for Sequencing DNA (Maxam and Gilbert Technique)	340
15.5.	Sequencing of RNA	342
15.6.	Gel Electrophoresis for Working Out Sequence of Bases in DNA Fragment	342
15.7.	Horizontal Ultrathin Gel Electrophoresis (HUGE)	347
15.8.	Automated Sequencing	348
15.9.	Polymerase Chain Reaction (PCR)	350
15.9.1.	PCR—Past and Present	351
15.9.2.	PCR Cycles	351
15.9.2.1.	The Cycling Reactions	351
15.9.2.2.	Designing/Selection of Primers	353
15.9.2.3.	Potential Problems in PCR	357
15.9.2.4.	Long PCR, Kinetic PCR, in situ PCR	357
15.9.2.5.	AP-PCR or RAPD PCR	358
15.9.2.6.	“Touchdown” PCR	359
15.9.2.7.	RAPD PCR	359
15.9.2.8.	Applications	360

DNA : The Basic Molecule of 'Gene'

1.1. INTRODUCTION

Major molecules that constitute the living cells (biomolecules) are classified according to their physical and chemical properties. Among these nucleotides are most versatile. In comparison to other classes of molecules—amino acids, carbohydrates, lipids, etc., nucleotides are notable for their involvement in the reactions which are very essential for maintenance and propagation of life. The nucleotides are known to participate in the process of energy transfer. When in polymeric forms these constitute nucleic acids which have main role in storing and encoding the genetic informations. These molecules (nucleotides and nucleic acids both) have been now, for about two decades, known to play catalytic roles in the cells. There are varied and essential functions of life which nucleic acid performs in the cell and no other molecule is known to perform such varied functions.

The nucleotides show considerable diversity in their structure. There are eight common varieties of nucleotides which are composed of nitrogenous base linked to a sugar (Base + sugar = nucleoside) to which at least one phosphate group is also attached (Base + sugar + phosphate = nucleotide). The bases of the nucleotides are planer, aromatic and heterocyclic molecules which are structural derivatives of either purine or pyrimidine although *in vivo* these are never synthesized from either of these organic molecules.

DNA molecule stores the genetic information and it is interesting to

note that this information is stored in DNA molecules as a sequence of nucleotides. The genetic engineering deals with *in vitro* manipulation of these sequences through recombinant DNA technology. It is, therefore, very important to brush up our understanding of structure and organization of nucleotides in DNA molecules.

Most common purines are adenine (A) and guanine (G) and the major pyrimidines are cytosine (C), thymine (T) and uracil (U). The purines bind to 5-C sugar (Pentose via their N-1 atoms while the pyrimidines bind to 5-C sugar via N-1 atoms.

The ribonucleotides constitute (as polymer) the RNA wherein pentose is ribose, the deoxyribonucleotides constitute DNA wherein pentose is 2-deoxyribose (i.e. carbon position 2 lacks a hydroxyl group). The primed numbers always indicate carbon position in the sugars. The carbon position in the bases is represented as unprimed number. The usual structures of ribonucleotides and deoxyribonucleotides are represented in Fig. 1.1.

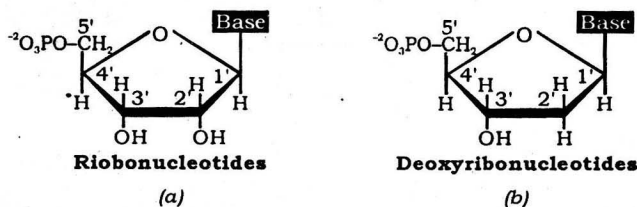
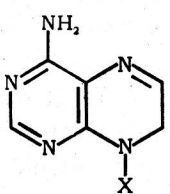
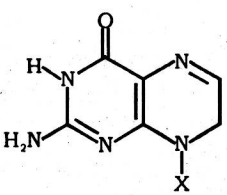
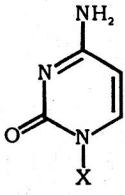
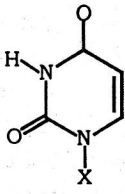
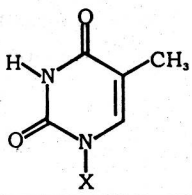


Fig. 1.1: Chemical structure of (a) ribonucleotides and (b) deoxyribonucleotides. The purine or pyrimidine base is linked to C-1' of the pentose and to the last one phosphate is attached.

Whereas adenine, guanine and cytosine are common to both RNA and DNA, the thymine is present only in DNA and instead of thymine uracil is present in RNA. The bulk of the nucleotides in the cell are found in a polymeric forms as either DNA or RNA whose primary functions are storage or transfer of information. However free nucleotides or their derivatives may also be present in the cells and may perform variety of metabolic functions which may not be related to the management of genetic information. One of the best known example of such derivatives is adenosine triphosphate (ATP - the energy currency of the cell). Some other such molecules are adenosine diphosphate (ADP), flavin adenine dinucleotide (FAD) adenosine linked to riboflavin (Vit. B₂) via two phosphate groups, nicotinamide adenine dinucleotide (NAD⁺) or its related compound, nicotinamide adenine dinucleotide phosphate (NADP⁺) and coenzyme-A (CoA). It is very well established now that these molecules have remained as key players in the metabolic field (even during prebiotic-evolutionary succession) while other structures and processes involving nucleotides have

grown up around them (Table 1.1).

TABLE 1.1: Names and Abbreviations of Nucleic Acid Bases, Nucleosides and Nucleotides

<i>Base Formula</i>	<i>Base (X = H)</i>	<i>Nucleoside (X = ribose^a)</i>	<i>Nucleotide (X = ribose phosphate)</i>
	Adenine Ade A	Adenosine Ado A	Adenylic acid Adenosine monophosphate AMP
	Guanine Gua G	Guanosine Guo G	Guanylic acid Guanosine monophosphate GMP
	Cytosine Cyt C	Cytidine Cyd C	Cytidylic acid Cytidine monophosphate CMP
	Uracil Ura U	Uridine Urd U	Uridylic acid Uridine monophosphate UMP
	Thymine Thy T	Deoxythymidine dThd dT	Deoxythymidylic acid Deoxythymidine monophosphate dTMP

1.2. STRUCTURE OF NUCLEIC ACIDS

As we have known the nucleic acids are polymers of nucleotides in the form of either DNA or RNA. In the polymers the neighbouring units of

sugars (ribose or deoxyribose) are bridged together in between position 3' (3 prime) and 5' (5 prime). As phosphate of these polynucleotides are acidic, so at physiological pH, nucleic acids are poly-anions.

The bond between two sugar units is known as phosphate diester bond as the phosphate is esterified by two sugar units. Each nucleotide that enters into the polynucleotide is known as nucleotide residue. In a polymeric molecule, the terminal sugar residue either at C5' or C3' is not linked to another nucleotide. The terminal residue whose C5' end is not linked to another nucleotide is called 5' end of the strand while the terminal residue whose C3' is not linked to another nucleotide is termed as 3' end of the strand. By property of polymerised units it may be **monomer** (single nucleotide) **dimer**, **trimer**, **tetramer** (2, 3, 4 units polymerised respectively) and so on through oligomer (a few units polymerised). As the size of the polymer increases the physical properties such as charge on the molecules, solubility and several other characteristics may change. When a polymer has similar repeating units, it is called homopolymer but when molecule has a series of non-identical units/residues it is called heteropolymer and has a property of containing some specific information in the form of its sequence of residues.

1.3. BASE COMPOSITION OF DNA : CHARGAFF'S RULE

In the late 1940s Erwin Chargaff devised the first reliable quantitative method for compositional analysis of DNA. It was demonstrated that DNA molecules have equal numbers of adenine and thymine. Similarly guanine and cytosine are present in equal numbers. It was also shown later that DNA base composition varies widely among different organisms. In bacteria for example it ranges from about 25% to 75% G+C, however, it is more or less constant among related species. In mammals G+C ranges from 39-46%. The significance of Chargaff's rule was not appreciated well at that time but with the establishment of double stranded nature of DNA, this rule has strengthened the concept and structural basis of the molecule. It is important here to remember that RNA being single stranded structure does not obey this rule.

1.3.1. Structure of DNA

After the initial experiments by Griffith (1920) on process of transformation in strains of *Diplococcus pneumoniae* (= *Streptococcus pneumoniae*) Avery, McLeod and McCarty indicated that the nature of the transforming factors was DNA. Later Hershey and Chase (1952) conclusively established that DNA was the genetic material through which the genetic

information is passed on from one generation to the other generation.

These discoveries exerted lot of pressure on scientists to establish the structure of DNA and lot of competition was set on them to become first to discover its chemical and physical structure. People at this time also knew that nucleotides were the building blocks of DNA and that these were linked by dehydration synthesis to form DNA but the exact structure/arrangement was unknown.

In the early 1950s Rosalind Franklin carried out experiments which involved bouncing X-rays off the crystals of various substances (a process which is called X-ray crystallography) including DNA, then exposing photographic film to bouncing X-rays. The scatter pattern of X-rays bouncing off the crystals of DNA is presented as a landmark photograph and is depicted below in Fig. 1.2.

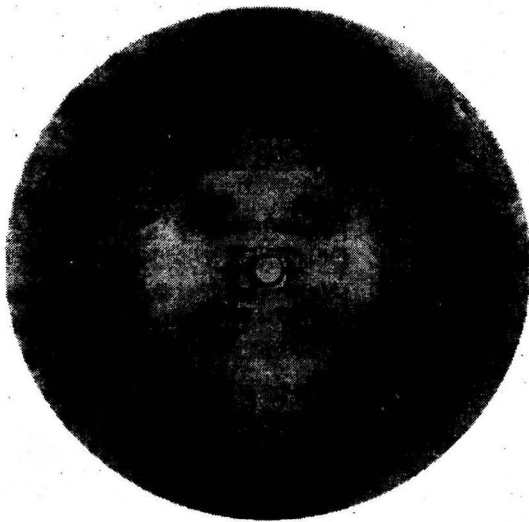


Fig. 1.2: X-ray diffraction pattern of DNA crystals as obtained by Franklin and Wilkins.

Other people like Linus Pauling were also attempting to figure out the structure of DNA. James Watson while working with Francis Crick were shown Franklin's photographs of DNA X-ray crystallography. From her pictures, and basing their imagination on Chargaff's data on DNA composition Watson and Crick published a paper in 1953 in which they proposed and described a hypothetical structure for DNA. They suggested that DNA was organised into a double spiral or double helix. Some of the major characteristics of DNA as suggested by them are

1. DNA is a double helix. Two Strands of the helix wound around each other to form a right handed helix.