Biotechnology of Nucleic Acids Medicines as Gene Therapeutics and Their Drug Complexes

Loutfy H Madkour*

Chemistry Department, Faculty of Science and Arts, Baljarashi, Al Baha University, Baljarashi 65635, Saudi Arabia

***Corresponding Author:** Loutfy H. Madkour Chemistry Department, Faculty of Science and Arts, Baljarashi, Al Baha University, Baljarashi 65635, Saudi Arabia.*

Abstract

Nucleic acids play an important role in cellular processes including cell division (DNA replication) and protein synthesis (transcription and translation). The deoxyribonucleic acid or DNA contains the genetic information necessary for the replication of the cell and its functioning after transcription of DNA into RNA and translation of RNA into proteins. These processes occur both in healthy cells, and in cancer cells, in which case they are targets for anti-cancer drugs. Two groups of medications concerning DNA: those, which protect DNA from alterations, due for example to radical reactions, causing cancers and those, which damage DNA of cancerous cells and microorganisms, while respecting if possible DNA of normal cells. Drugs, which damage nucleic acids, DNA and RNA, already constituted, inhibit cellular replication and function.

They are used as antineoplastic agents and as antibiotics. Lipid-based vesicles are a very promising approach to treat diseases such as cancer, chronic infections and auto-immunity. Modern drug encapsulation methods allow efficient packing of therapeutic substances inside liposomes, thereby reducing the systemic toxicity of the drugs. Gene therapy is a technique for correcting defective genes responsible for disease development, which may be classified into two types: somatic and germ line gene therapy. Nucleic acid-based molecules (deoxyribonucleic acid, complementary deoxyribonucleic acid, complete genes, ribonucleic acid, and oligonucleotides) are utilized as research tools within the broad borders of gene therapy and the emerging field of molecular medicine. "Nucleic acid medicine" is a next generation drug discovery technology with a completely different mechanism of action than traditional pharmaceutical products.

Keywords: *Nucleic acid medicine; DNA-based therapeutics; RNA-based therapeutics; Gene therapy technology; Drug Complexes*

Introduction

Most human diseases are complex multi-factorial diseases resulting from the combination of various genetic and environmental factors. At present, there is no drug really able to protect DNA; the only mean available is the decrease of the exposure to the risk factors such as sun, tobacco, irradiations, and toxic compounds. Specific targeting can enhance the therapeutic effect of the drugs through their

accumulation_at the diseased site. In the vaccine field, the integration of functional viral envelope proteins into liposomes has led to an antigen carrier and delivery system termed a virosome, a clinically proven vaccine platform for subunit vaccines with an excellent immunogenicity and tolerability profile.

DNA is a linear polymer, constituted of deoxyribonucleotides (deoxyadenosine, deoxyguanosine, deoxycytidine and deoxythymidine) bound by phosphodiester bonds. At pH = 7, the phosphodiester groups are ionized as anions and bind to elements such as magnesium. DNA exists as a double helix, or double strand, maintained between them by hydrogen bonds drawn up between two complementary bases, adenine and thymine by two hydrogen bonds, and guanine and cytosine by three hydrogen bonds.

The double helix of DNA undergoes supercoiling known as positive or negative [1] according to the direction of rotation. Under supercoiled form, replication and transcription of DNA are not possible. A specific and transitional separation of the strands, after their cut, is necessary. The enzymes which cut and then restore the bond on only one strand of the double helix are called topoisomerases I and those which act on the two strands are called topoisomerases II or DNA-gyrases. These enzymes can add or remove crossings. RNA is a polymer constituted of ribonucleotides in the form of one strand. The four bases of RNA are adenine, uracil, guanine and cytosine. DNA is a biological molecule having the capacity for self-repairing.

The damage affects generally only one strand of DNA and the intact strand is used as reference for the repair of the damaged strand. Moreover, the bases having bound an alkyl group can be de-alkylated by suicide enzymes which remove and bind the alkyl group. Finally there exist mechanisms of excision of damaged nucleotides. These mechanisms of repair are useful for repairing not desired lesions but become adverse when they antagonize the action of antineoplastic agents in the cancerous cells. One can schematically distinguish two groups of medications concerning DNA: those which protect DNA from alterations, due for example to radical reactions, causing cancers and those which damage DNA of cancerous cells and microorganisms, while respecting if possible DNA of normal cells. At present there is no drug really able to protect DNA; the only mean available is the decrease of the exposure to the risk factors such as sun, tobacco, and irradiations, toxic compounds.

This review is devoted to drugs which damage DNA. Drugs which damage nucleic acids, DNA and RNA, already antibiotics. The majority of antineoplastic agents damaging DNA are old drugs, not very specific because they do not respect healthy cells and have many adverse effects, but their use, generally in combination, gives appreciable results. The antibiotics which damage DNA have a sufficient specificity of action against the pathogenic microorganisms to be generally well tolerated.

Twenty years ago, the Human Genome Project was initiated aiming to uncover the genetic factors of human diseases and to develop new strategies [2] for diagnosis, treatment and prevention. The successful sequencing of the human genome and the following coordinated efforts, such as the HapMap project [3], genome-wide association studies [4] and the cancer genome projects [5], have resulted in the discovery of many disease-associated genes. However, our understanding of molecular mechanisms is still largely incomplete for the majority of diseases, which are multi-factorial diseases resulting from the combination of various genetic and environmental factors.

There must be inherent relationships among these factors for the etiology and pathogenesis, and they may be characterized by considering the molecular networks involving these factors. The analysis of network–disease associations, in addition to gene–disease associations, would better clarify the molecular mechanisms of diseases and help develop new drugs and treatments. In recent years increased attention has been focused on the ways in which drugs interact with biological systems, with the goal of understanding the toxic as well as chemotherapeutic effects of these small molecules [6].

In the cell many drugs, particularly those with planar chromophores, bind to nucleic acids. It is thought that this complex formation may be the first step in mutagenesis and possibly carcinogenesis. Nucleic acids with their evenly stacked base pairs and shallow and deep grooves are attractive targets for these molecules. Planar drugs can intercalate between base pairs, drawing them apart from their normal 3.4°A spacing to 6.8°A while bulkier drugs can fit into either groove, sometimes with minimal distortion of the structure.

Both procedures can profoundly influence the recognition properties of the nucleic acid. Some drugs such as actinomycin D, which is a powerful antibiotic, and daunomycin, a potent chemotherapeutic agent, exhibit both binding modes.

What is Deoxyribonucleic Acid (DNA)?

DNA, or deoxyribonucleic acid, is the hereditary material [7] in humans and almost all other organisms. Nearly every cell in a person's body has the same DNA. Most DNA is located in the cell nucleus (where it is called nuclear DNA), but a small amount of DNA can also be found in the mitochondria (where it is called mitochondrial DNA or mtDNA). The information in DNA is stored as a code [7] made up of four chemical bases: adenine (A), guanine (G), cytosine (C), and thymine (T).

Human DNA consists of about 3 billion bases, and more than 99 percent of those bases are the same in all people. The order, or sequence, of these bases determines the information available for building and maintaining an organism, similar to the way in which letters of the alphabet appear in a certain order to form words and sentences. DNA bases pair up with each other, A with T and C with G, to form units called base pairs. Each base is also attached to a sugar molecule and a phosphate molecule. Together, a base, sugar, and phosphate are called a nucleotide. Nucleotides are arranged in two long strands that form a spiral called a double helix. The structure of the double helix is somewhat like a ladder, with the base pairs forming the ladder's rungs and the sugar and phosphate molecules forming the vertical sidepieces of the ladder.

An important property of DNA is that it can replicate, or make copies of itself. Each strand of DNA in the double helix can serve as a pattern for duplicating the sequence of bases. This is critical when cells divide because each new cell needs to have an exact copy of the DNA present in the old cell.
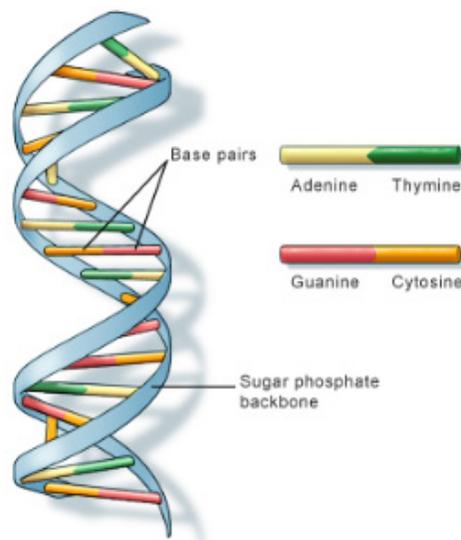
U.S. National Library of Medicine

Figure 1: DNA is a double helix formed by base pairs attached to a sugar-phosphate backbone.

For more information about DNA

The National Human Genome Research Institute fact sheet Deoxyribonucleic Acid (DNA) (https://www.genome.gov/25520880) introduces this molecule [7]. Information about the genetic code (https://geneed.nlm.nih.gov/topic_subtopic.php?tid=15&sid=19) and the structure of the DNA double helix (https://geneed.nlm.nih.gov/topic_subtopic.php?tid=15&sid=16) is available from GeneEd.
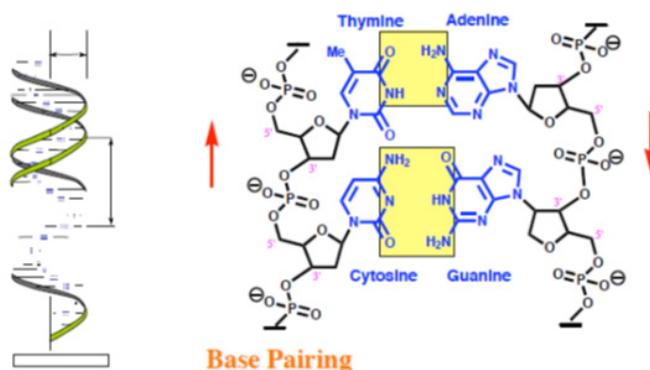
G-C base pairing involves 3 H-bonds

A-T base pairing involves 2 H-bonds

Figure 2: DNA bases pair up with each other.

The New Genetics, a publication of the National Institute of General Medical Sciences, discusses the structure of DNA and how it was discovered http://www.atdbio.com/content/14/Transcription-Translation-and-Replication (https:// publications.nigms.nih. gov/thenewgenetics/chapter1.html#c1). A basic explanation and illustration of DNA (https://askabiologist.asu.edu/dnashape-and-structure) can be found on Arizona State University's. The Virtual Genetics Education Centre, created by the University of Leicester, offers additional information on DNA, genes, and chromosomes (http://www2.le.ac.uk/projects/vgec/schoolscolleges/topics/dna-genes-chromosomes).

What is mitochondrial DNA?

Although most DNA is packaged in chromosomes within the nucleus, mitochondria also have a small amount of their own DNA. This genetic material is known as mitochondrial DNA or mtDNA. Mitochondria (Figure 3) are structures within cells that convert the energy from food into a form that cells can use. Each cell contains hundreds to thousands of mitochondria, which are located in the fluid that surrounds the nucleus (the cytoplasm).

Mitochondria

U.S. National Library of Medicine

U.S. National Library of Medicine

Figure 3: Mitochondria provide the cell's energy.

Mitochondria produce energy through a process called oxidative phosphorylation. This process uses oxygen and simple sugars to create adenosine triphosphate (ATP), the cell's main energy source. A set of enzyme complexes, designated as complexes I-V, carry out

oxidative phosphorylation within mitochondria. In addition to energy production, mitochondria play a role in several other cellular activities. For example, mitochondria help regulate the self-destruction of cells (apoptosis). They are also necessary for the production of substances such as cholesterol and heme (a component of hemoglobin, the molecule that carries oxygen in the blood). Mitochondrial DNA contains 37 genes, all of which are essential for normal mitochondrial function. Thirteen of these genes provide instructions for making enzymes involved in oxidative phosphorylation. The remaining genes provide instructions for making molecules called transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs), which are chemical cousins of DNA. These types of RNA help assemble protein building blocks (amino acids) into functioning proteins.

For more information about mitochondria and mitochondrial DNA

Molecular Expressions, a web site from the Florida State University Research Foundation, offers an illustrated introduction to mitochondria and mitochondrial DNA (http://micro.magnet.fsu.edu/cells/mitochondria/mitochondria.html). An overview of mitochondrial DNA (http://neuromuscular.wustl.edu/mitosyn.html#general) is available from the Neuromuscular Disease Center at Washington University.

Tertiary Structure [8]

- Double helix coils into a 3D shape - supercoiling
- Double helix has to unravel during replication
- Unravelling leads to strain
- Relieved by enzyme catalysed cutting and repair of DNA chain
- Important to the activity of the quinolone and fluoroquinolone antibacterial agents which act as enzyme inhibitors.

Ribonucleic Acid (RNA)

Primary structure

- Similar to DNA with the following exceptions
- Ribose is used instead of deoxyribose
- Uracil is used rather than thymine
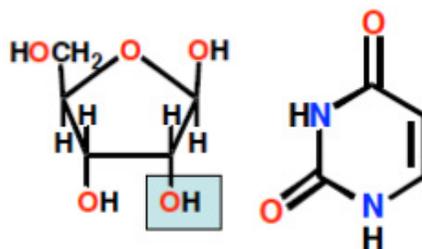
Ribose Uracil

Secondary structure

- Single stranded
- Some regions of helical secondary structure exist due to base pairing within the same strand.
- Adenine pairs to uracil; guanine pairs to cytosine.

Tertiary structure

- Three types of RNA are involved in protein synthesis:
- Messenger RNA (mRNA)
- Relays the code for a protein from DNA to the protein production site

- Transfer RNA (tRNA)
- The adapter unit linking the triplet code on mRNA to specific amino acids
- Ribosomal RNA (rRNA)
- Present in ribosomes (the production site for protein synthesis).
- Important both structurally and catalytically
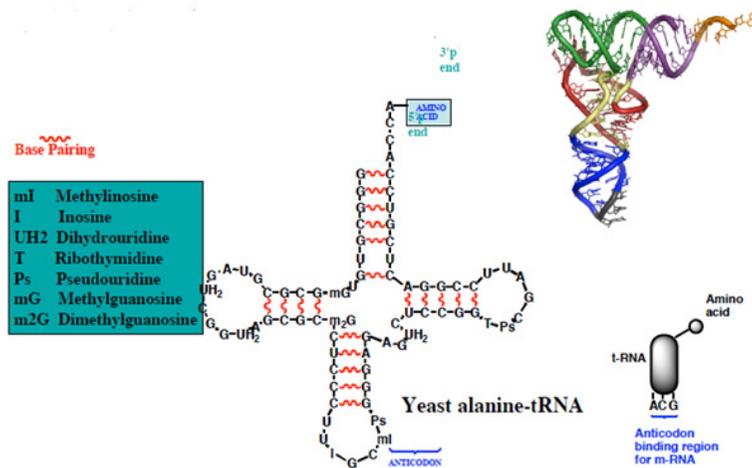
Anticodon-the 3 bases are specific for the attached amino acid-base pair to the complementary triplet code on m-RNA (the codon)

Figure 4: *Production site for protein synthesis.*

Transcription, Translation and Replication

DNA, RNA and protein synthesis

The genetic material is stored in the form of DNA in most organisms. In humans, the nucleus of each cell contains 3×10^9 base pairs of DNA distributed over 23 pairs of chromosomes, and each cell has two copies of the genetic material. This is known collectively as the human genome. The human genome contains around 30 000 genes, each of which codes for one protein.

Large stretches of DNA in the human genome are transcribed but do not code for proteins. These regions are called *introns* and make up around 95% of the genome. The nucleotide sequence of the human genome is now known to a reasonable degree of accuracy but we do not yet understand why so much of it is non-coding. Some of this non-coding DNA controls gene expression but the purpose of much of it is not yet understood. This is a fascinating subject that is certain to advance rapidly over the next few years. The *Central Dogma of Molecular Biology* states that DNA makes RNA makes proteins (Figure 5).
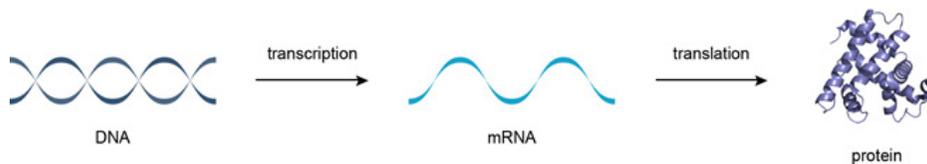

Figure 5: *The Central Dogma of Molecular Biology: DNA makes RNA makes proteins.*

The process by which DNA is copied to RNA is called transcription, and that by which RNA is used to produce proteins is called translation.

Mistakes in DNA replication

DNA replication is not perfect. Errors occur in DNA replication, when the incorrect base is incorporated into the growing DNA strand. This leads to *mismatched* base pairs, or *mispairs*. DNA polymerases have proofreading activity, and a DNA repair enzymes have evolved to correct these mistakes. Occasionally, mispairs survive and are incorporated into the genome in the next round of replication. These mutations may have no consequence, they may result in the death of the organism, they may result in a genetic disease or cancer; or they may give the organism a competitive advantage over its neighbours, which leads to evolution by natural selection.

Transcription

Transcription is the process by which DNA is copied (*transcribed*) to mRNA, which carries the information needed for protein synthesis. Transcription takes place in two broad steps. First, pre-messenger RNA is formed, with the involvement of RNA polymerase enzymes. The process relies on Watson-Crick base pairing, and the resultant single strand of RNA is the reverse-complement of the original DNA sequence. The pre-messenger RNA is then "edited" to produce the desired mRNA molecule in a process called RNA *splicing*.

Formation of pre-messenger RNA

The mechanism of transcription has parallels in that of DNA replication. As with DNA replication, partial unwinding of the double helix must occur before transcription can take place, and it is the RNA polymerase enzymes that catalyze this process.

Unlike DNA replication, in which both strands are copied, only one strand is transcribed. The strand that contains the gene is called the *sense* strand, while the complementary strand is the *antisense* strand. The mRNA produced in transcription is a copy of the sense strand, but it is the antisense strand that is transcribed. Ribonucleotide triphosphates (NTPs) align along the antisense DNA strand, with Watson-Crick base pairing (A pairs with U). RNA polymerase joins the ribonucleotides together to form a pre-messenger RNA molecule that is complementary to a region of the antisense DNA strand. Transcription ends when the RNA polymerase enzyme reaches a triplet of bases that is read as a "stop" signal. The DNA molecule re-winds to re-form the double helix.
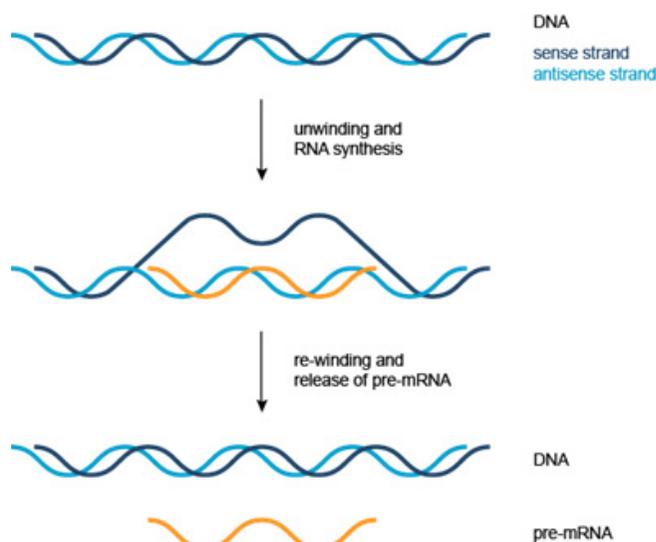
Figure 6: Transcription Simplified representation of the formation of pre-messenger RNA (orange) from double-stranded DNA (blue) in transcription.

Reverse transcription

In reverse transcription, RNA is "reverse transcribed" into DNA. This process, catalyzed by reverse transcriptase enzymes, allows retroviruses, including the human immunodeficiency virus (HIV), to use RNA as their genetic material. Reverse transcriptase enzymes have also found applications in biotechnology, allowing scientists to convert RNA to DNA for techniques such as PCR.

Translation

The mRNA formed in transcription is transported out of the nucleus, into the cytoplasm, to the ribosome (the cell's protein synthesis factory). Here, it directs protein synthesis. Messenger RNA is not directly involved in protein synthesis – transfer RNA (tRNA) is required for this. The process by which mRNA directs protein synthesis with the assistance of tRNA is called *translation*.

The ribosome is a very large complex of RNA and protein molecules. Each three-base stretch of mRNA (triplet) is known as a *codon*, and one codon contains the information for a specific amino acid. As the mRNA passes through the ribosome, each codon interacts with the anticodon of a specific transfer RNA (tRNA) molecule by Watson-Crick base pairing. This tRNA molecule carries an amino acid at its 3'-terminus, which is incorporated into the growing protein chain. The tRNA is then expelled from the ribosome. Figures 7 and 8 shows the steps involved in protein synthesis.
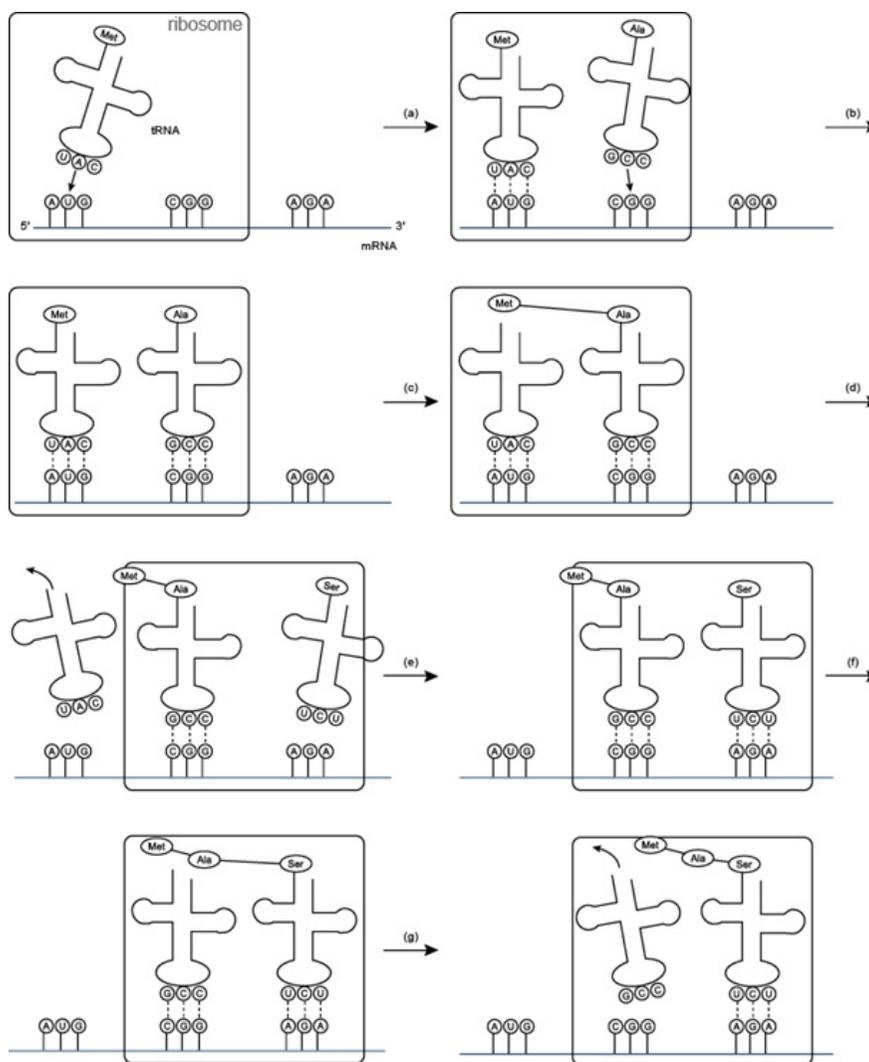
Figure 7: *Translation (a) and (b) t RNA molecules bind to the two binding sites of the ribosome, and by hydrogen bonding to the mRNA; (c) a peptide bond forms between the two amino acids to make a dipeptide, while the t RNA molecule is left uncharged; (d) the uncharged t RNA molecule leaves the ribosome, while the ribosome moves one codon to the right (the dipeptide is translocated from one binding site to the other); (e) another t RNA molecule binds; (f) a peptide bond forms between the two amino acids to make a tripeptide; (g) the uncharged t RNA molecule leaves the ribosome.*
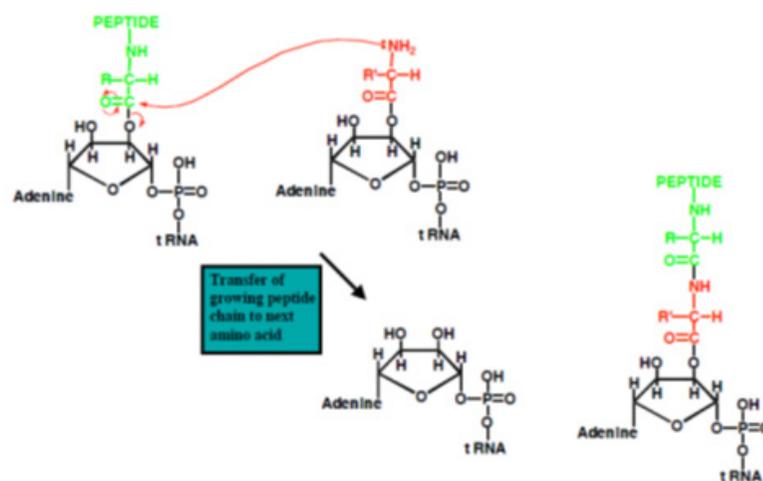
Figure 8: *Translation-protein synthesis.*

**Translation-protein synthesis: Overview**
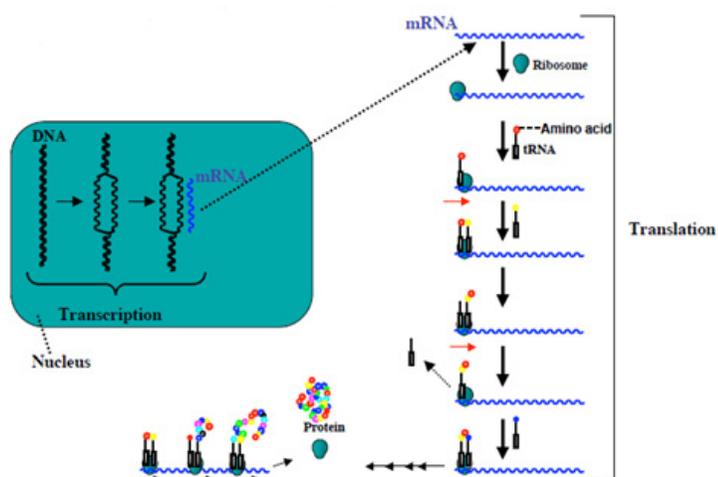
Figure 9: *Overview protein synthesis.*

Drugs Acting on DNA

Until the 1980s, most of our knowledge about drugs [9], drug mechanisms and drug receptors could fit in a few encyclopedic books and a couple dozen schematic figures. However, with the recent explosion in biological and chemical knowledge, this is no longer the case. There is simply too much data (images, models, structures and sequences) from too many sources. Unfortunately, most of this information still resides in textbooks or print journals. The limited drug or drug receptor data that is electronically available is either inaccessible (except through expensive subscriptions), inadequate or widely scattered among many different public databases. This state of affairs largely reflects the 'two solitudes' of cheminformatics and bioinformatics. Neither discipline has really tried to integrate with the other. As a consequence, the wealth of electronic sequence/structure data that exists today has never been well linked to the enormous body of drug or chemical knowledge that has accumulated over the past half century.

Recently, some notable efforts have been made to partially overcome this 'informatics gap'. The Therapeutic Target Database or TTD is one such example [10]. This very useful web-based resource contains linked lists of names for >1100 small molecule drugs

and drug targets (i.e. proteins). In addition to the TTD, a number of more comprehensive small molecule databases have also emerged including KEGG [11], ChEBI [12] and PubChem (http://pubchem.ncbi.nlm.nih.gov/).

Each contains tens of thousands of chemical entries-including hundreds of small molecule drugs. All three databases provide names, synonyms, images, structure files and hyperlinks to other databases. Furthermore, both KEGG and PubChem support structure similarity searches. Unfortunately, these databases were not specifically designed to be drug databases, and so they do not provide specific pharmaceutical information or links to specific drug targets (i.e. sequences). Furthermore, because these databases were designed to be synoptic (containing < 15 fields per compound entry) they do not provide a comprehensive molecular summary of any given drug or its corresponding protein target.

Intercalating agents

Mechanism of action
- Contain planar aromatic or heteroaromatic ring systems.
- Planar systems slip between the layers of nucleic acid pairs and disrupt the shape of the helix.
- Preference is often shown for the minor or major groove.
- Intercalation prevents replication and transcription.
- Intercalation inhibits topoisomerase.

Examples

Dactinomycin Doxorubicin (Adriamycin)
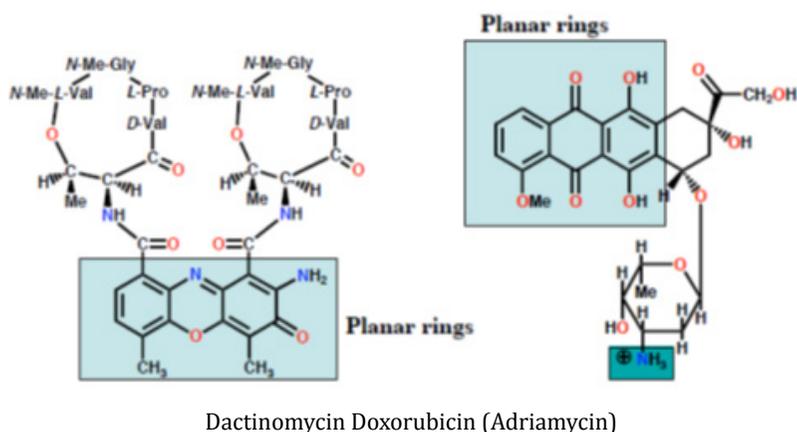
Extra binding to sugar phosphate Extra binding to sugar backbone by cyclic peptide phosphate backbone by NH_3

Figure 10: Mechanism of action of intercalating agents.

Alkylating agents
- Contain highly electrophilic groups
- Form covalent bonds to nucleophilic groups in DNA (e.g. 7-N of guanine)
- Prevent replication and transcription
- Useful anti-tumour agents
- Toxic side effects (e.g. alkylation of proteins)

Examples

Mechlorethamine (nitrogen mustard)
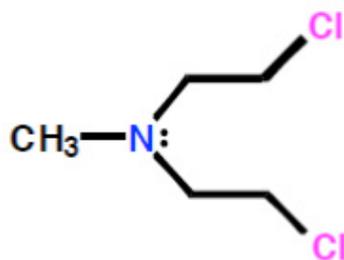
Cross linking
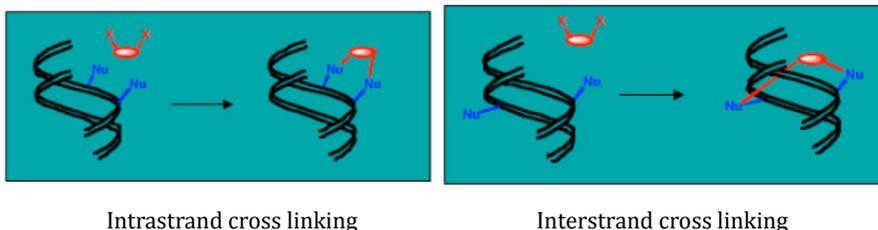
Intrastrand cross linking   Interstrand cross linking

Mechanism of action

Figure 11: Mechanism of action mechlorethamine as alkylating agents for drugs acting on DNA.

Mechlorethamine analogues
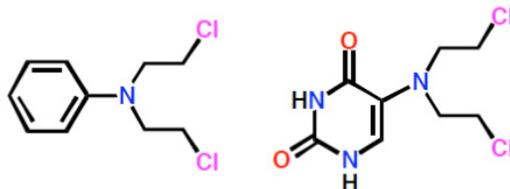
**Aromatic ring - e withdrawing effect
N is less nucleophilic
Less reactive alkylating agent
Selective for stronger nucleophiles
(e.g. guanine)**

**Uracil mustard
Used vs leukemia
Attached to a nucleic acid building
block
Concentrated in fast growing cells
(tumours)
Some selectivity**

Cisplatin Mitomycin C
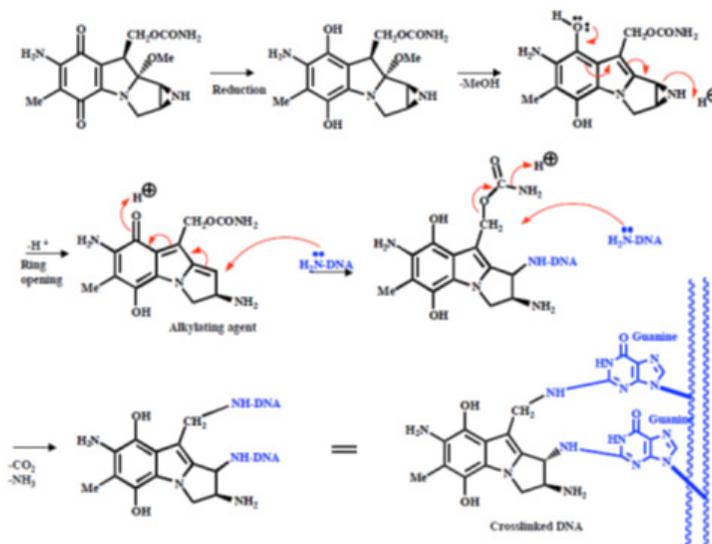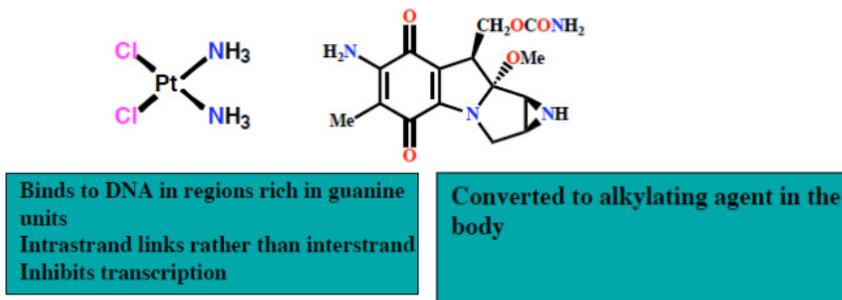
Figure 12: Mechanism of alkylating agent and crosslinked DNA.

Chain cutters

- Abstracts H from DNA to generate radicals
- Radicals react with oxygen resulting in chain cutting
- Bleomycin also inhibits repair enzymes.
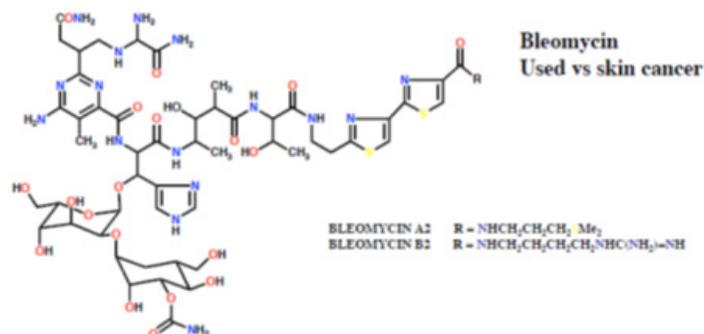
Figure 13: Representing chain cutting for drugs acting on DNA.

Drugs Acting On RNA

Drugs acting on r RNA: Antibiotics

Chlortetracycline Erythromycin (Aureomycin)
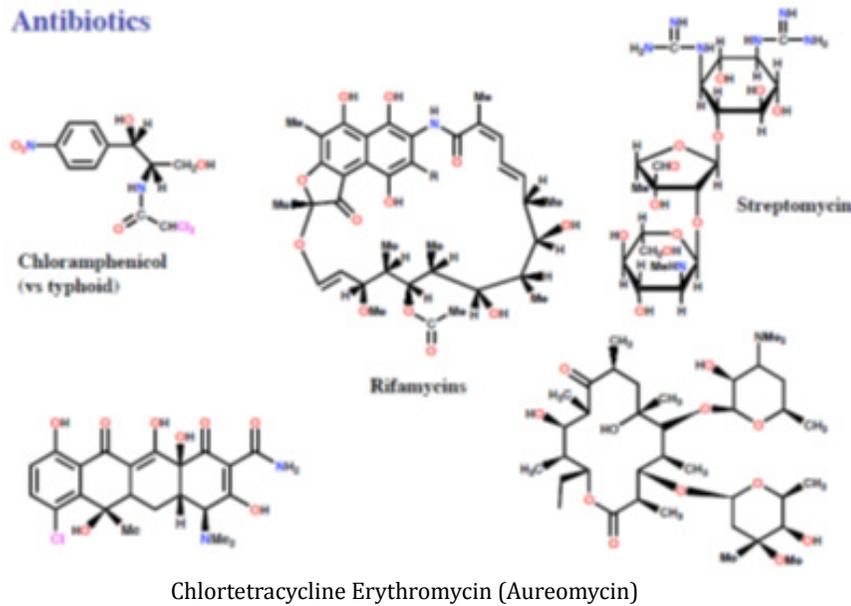
Figure 14: Representing Drugs acting on rRNA: Antibiotics.

Drugs acting on mRNA: Antisense RNA Therapy
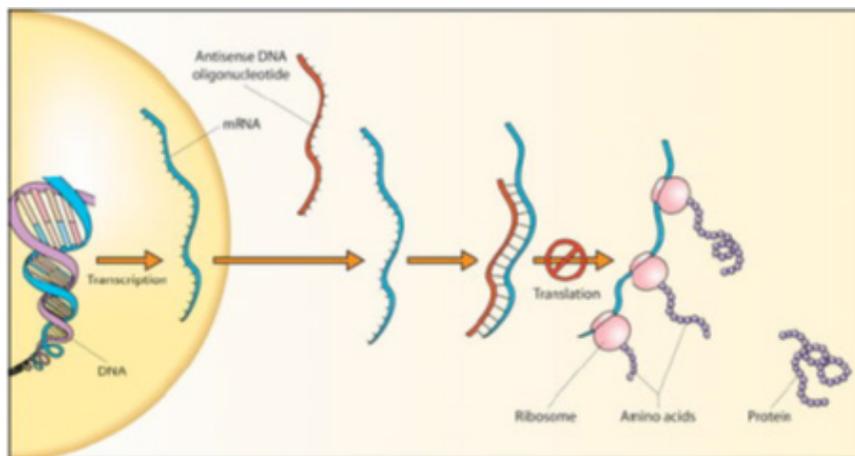
Figure 15: Antisense RNA Therapy.

Advantages

- Same effect as an enzyme inhibitor or receptor antagonist
- Highly specific where the oligonucleotide is 17 nucleotides or more
- Smaller dose levels required compared to inhibitors or antagonists
- Potentially less side effects

Disadvantages

• 'Exposed' sections of mRNA must be targeted

• Instability and polarity of oligonucleotides (pharmacokinetics)

• Short lifetime of oligonucleotides and poor absorption across cell membranes.

Drugs acting on or through RNA: Small Interfering RNA-siRNA
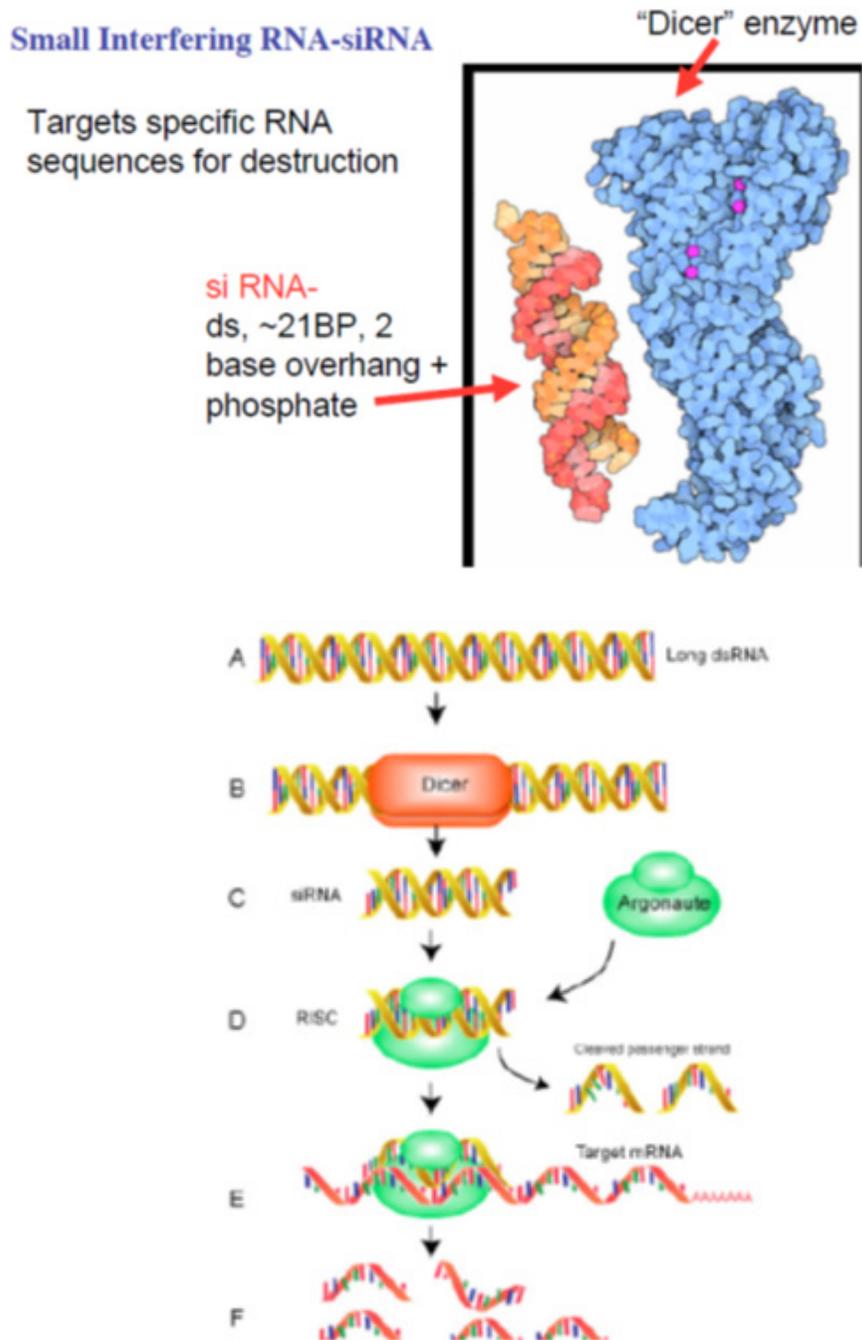
Figure 16: Representation of Small Interfering RNA-siRNA.

Drugs related to nucleic acid building blocks

Examples: Antiviral agents
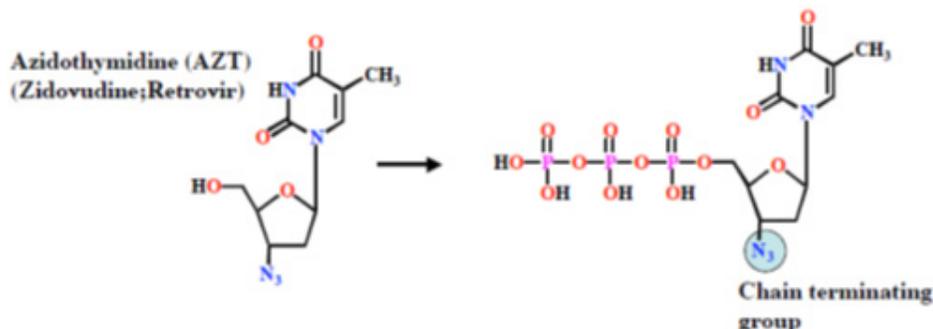
- Enzyme inhibitor
- AZT is phosphorylated to a triphosphate in the body
- Triphosphate has two mechanisms of action
  - inhibits a viral enzyme (reverse transcriptase)
  - added to growing DNA chain and acts as chain terminator
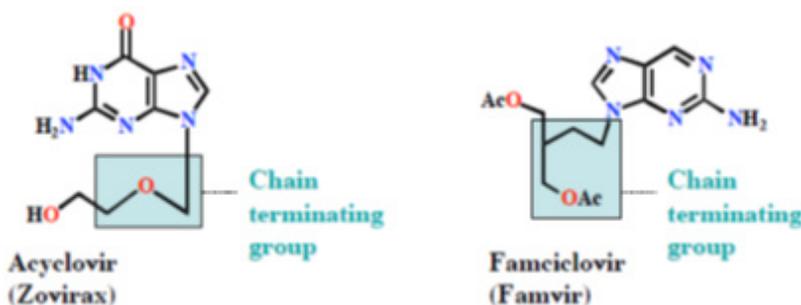
Notes:

Same mechanisms of action as AZT Used vs herpes simplex and shingles.

Nucleic Acid-Drug Interactions

Small molecules may interact with or bind with DNA. Such compounds can be classified by their mechanism of action. The main classes of DNA binding molecules are [13]:

- Groove binders that sit in the minor groove;
- Intercalators that sandwich between base pairs;
- Alkylators that can chemically react with DNA, resulting in DNA alkylation; and
- DNA cleavage agents that have the ability to break DNA chains.
- Each of these classes of molecules has a different structure and interacts with DNA in a different way.

Groove binders

Minor groove binding molecules are usually constructed of a series of heterocyclic or aromatic hydrocarbon rings that possess rotational freedom. This allows the molecule to fit into the minor groove, with displacement of water.

Distamycin and netropsin

Distamycin and netropsin (Figure 17) are natural products possessing amido groups and, respectively, three and two N-methylpyrrole rings (distamycin can be denoted PyPyPy, and netropsin PyPy).

Distamycin   Netropsin
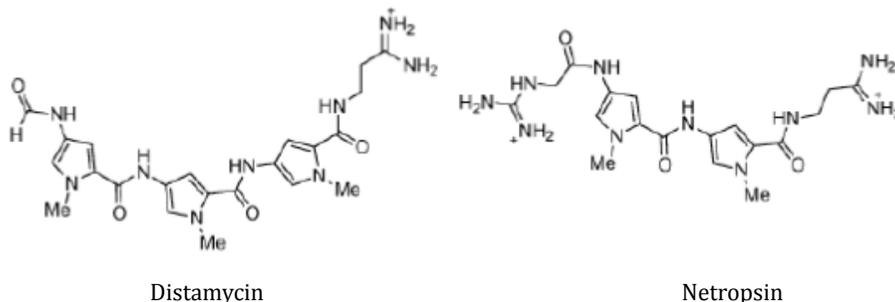
Figure 17: *Distamycin and netropsin Structures of the naturally-occurring minor groove binders distamycin and netropsin.*

Distamycin and netropsin interact with AT-rich regions of DNA in the minor groove by forming hydrogen bonding and hydrophobic interactions (Figure 18). The terminal amidine group of the small molecule is basic, and serves to attract the drug molecule to the negatively charged DNA phosphodiester backbone. The 2- amino group of guanine prevents distamycin from binding to the minor groove of G·C base pairs by steric hindrance, thus conferring AT-selectivity on the drug molecule.
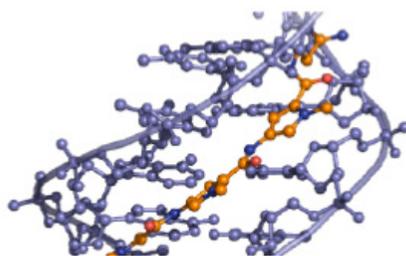
Figure 18: *Distamycin DNA binding View from the three-dimensional structure of a complex between distamycin (orange) and a DNA duplex, showing the binding of distamycin in the minor groove.*

Lexitropsins

A series of dimers and trimers of distamycin and netropsin have been synthesized and studied in an attempt to increase the DNA binding region from 3 base pairs (monomeric drug) to 10 base pairs or more. These semi-synthetic compounds have been named Lexitropsins. As well as pyrrole rings (Py), some lexitropsins also incorporate imidazole (Im) rings. The ability to recognize specific DNA sequences of more than 10 base pairs would give rise to a powerful tool in molecular biology and antisense/antigene therapeutics.

Dervan polyamides

Unfortunately, simple oligomeric compounds like distamycin and netropsin do not have the ideal crescent shape to wrap around the minor groove of DNA, and they fail to recognize longer stretches of DNA. Dervan took this approach further in synthesizing a series of oligomeric "hairpin" polyamide molecules containing pyrrole and imidazole ring systems that are able to bind side-by-side in the minor groove of DNA with high affinity and in a sequence-specific manner. These molecules can be prepared by solid-phase methods.

Intercalators

Certain flat aromatic or heteroaromatic molecules can slide between the base pairs of DNA (intercalate) and stabilize the duplex without disrupting base pairing. Intercalation has the effect of lengthening the duplex by around 3 Å per bound drug molecule, causes unwinding of DNA, and prevents replication and transcription by interfering with the action of topoisomerases. The degree of unwinding depends on the structure of the intercalating molecule and the site of intercalation. The tight ternary complex formed between the intercalated drug, the DNA and the topoisomerase is lethal to proliferating cells, so intercalators are often more toxic to cancer cells than to normal cells.

Acridines

Amsacrine
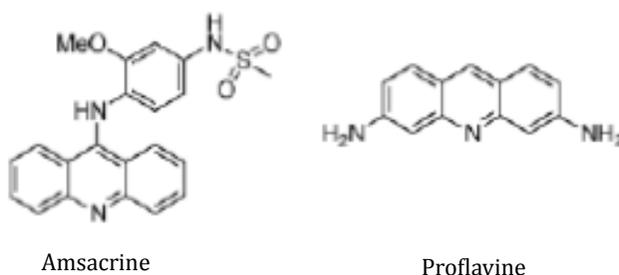
Proflavine

Figure 19: *Amsacrine and proflavine Structures of the acridines amsacrine and proflavine, which intercalate between DNA base pairs.*

Acridines originated from the aniline dye industry and have been used as anti-malarial and antibacterial drugs. Amsacrine (Figure 19) is used in the treatment of leukemia and proflavine (Figure 19) was used in the Second World War to treat wounds. Proflavine contains amino groups that interact ionically with the negatively charged phosphates groups on DNA, whilst the aromatic ring system intercalates.

Polypeptides

The Actinomycins (Figure20) are polypeptide antibiotics isolated from Streptomyces strains.
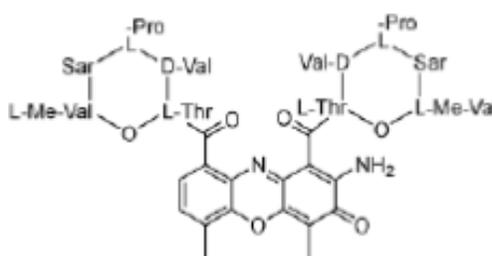
Figure 20: *Actinomycin D Structure of actinomyin D (dactinomycin), a polypeptide DNA-binding molecule.*

Actinomycins inhibit both DNA synthesis and RNA synthesis by blocking chain elongation. They interact with G·C base pairs as they require the 2-amino group of guanine for binding. The phenoxazone ring slides into the double helix and intercalates, while the pentapeptide side chains interact with the DNA minor groove by forming hydrogen bonding and hydrophobic interactions. The result of these two mechanisms of interaction between small molecule and DNA (intercalation and minor-groove binding) is a very stable complex (Figure 21).
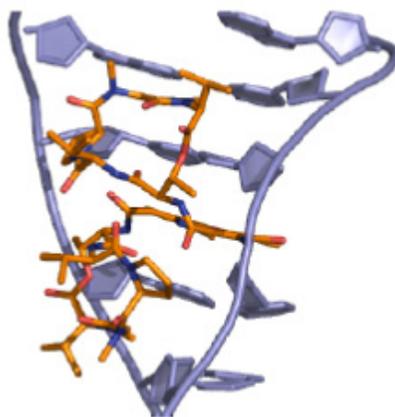
Figure 21: Actinomycin DNA binding View from the three-dimensional structure
of a complex between actinomycin D (orange) and a DNA duplex, showing the
intercalation of actinomycin D in double-stranded DNA.

Anthracyclines

Doxorubicin (adriamycin) and daunorubicin (daunomycin), isolated from *Streptomyces* strains, are important examples of anthracycline antitumour antibiotics. Both possess an amino group on the sugar which, when protonated, forms an ionic interaction with the negatively charged DNA phosphate backbone. This bond helps to hold the molecule in place, allowing the planar aromatic ring system to slide into the double helix. Although doxorubicin and daunorubicin differ by only one hydroxyl group, they have different activities: daunorubicin is active only against leukaemias, but doxorubicin is active against leukaemias and a wide range of solid tumours.

Doxorubicin (Adriamycin)   Daunorubicin (Daunomycin)
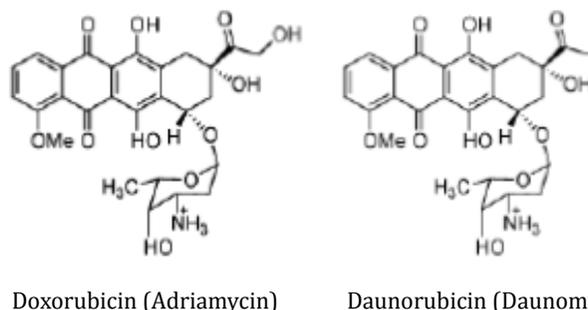
*Figure 22: Doxorubicin and daunorubicin Structures of the anthracyclines
doxorubicin (adriamycin) and daunorubicin (daunomycin).*

Alkylators

Alkylators are strongly electrophilic compounds that react chemically with nucleophilic groups on DNA to form covalent bonds. The resulting DNA adducts are irreversible inhibitors of transcription and translation. Nucleophilic substitution reactions at the DNA bases occur by both S_N1 and S_N2 mechanisms. The most reactive sites are those that are both nucleophilic and exposed in the grooves of the DNA duplex. The N (7) atom of guanine and the N (3) atom of adenine fulfil both of these criteria. Simple nucleophiles, for example ethyleneimines and methane sulfonates, tend to react via a S_N2 mechanism, whereas the nitrogen mustards can form aziridinium ions that react via an S_N1 mechanism.
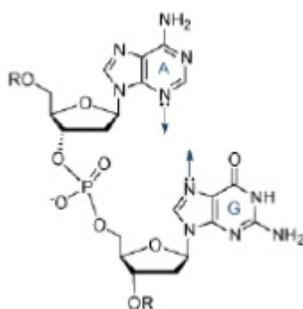
Figure 23: *Nucleophilicity of adenine and guanine The N (7)*
atom of guanine and the N (3) atom of adenine are nucleophilic
and are exposed in the grooves of the DNA duplex.

There are several different classes of DNA alkylators, including nitrogen mustards, ethyleneimines, methanesulfonates, nitro-soureas, triazenes and cis platinum complexes.

Mustards

Sulfur mustard (Figure 24) is a highly toxic nerve gas that was used during the First World War. Although too toxic for use in cancer therapy, it led to the development of a series of compounds known as nitrogen mustards (Figure 24) that signalled the beginning of modern cancer chemotherapy. Mechlorethamine (Figure 24) was the first of these compounds to be used in the treatment of advanced Hodgkin's disease.

Sulfur mustard   nitrogen mustard   mechlorethamine
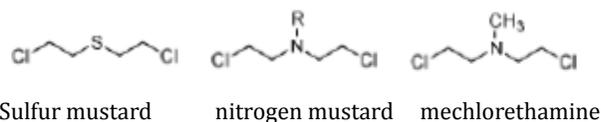
Figure 24: *Sulfur and nitrogen mustards General structures of*
sulfur mustards and nitrogen mustards; and the structure of mech-
lorethamine, a nitrogen mustard.

The mechanism of action of the mustards begins with the formation of an electrophilic aziridinium ion by displacement of chloride. The aziridinium ion is readily attacked by the nucleophilic DNA bases and the reaction sequence can occur a second time to cross-link two strands of DNA (Figure 25).
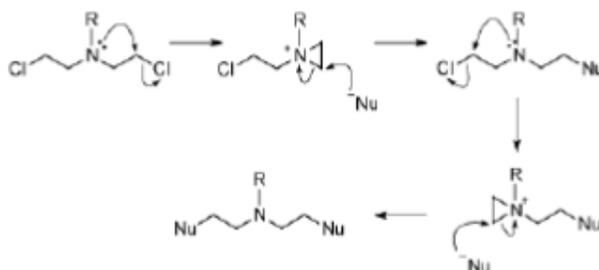
Figure 25: Mustard mechanism Mechanism of action of
mustards, alkylating agents that cross-link DNA strands.

Mechlorethamine reacts principally with the N (7) atom of guanine, and quaternization of this nitrogen atom leads to scission of the N-glycosidic bond (Figure26).
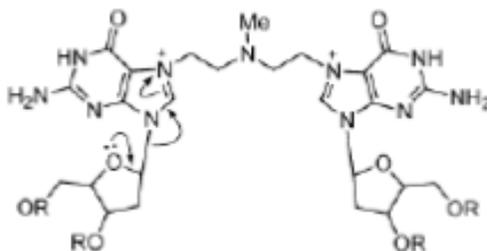
Figure 26: *Mechlorethamine mechanism Mechanism of action*
of mechlorethamine, a nitrogen mustard alkylating agent.

Mechlorethamine is a reactive molecule. It has a lifetime of only minutes in the body and is rapidly hydrolysed. Substitution of the methyl group with an electronwithdrawing aryl group reduces the nucleophilicity of nitrogen, slowing down the rate of aziridinium ion formation and therefore stabilising the compound. However, such compounds are not sufficiently water soluble for intravenous administration. This problem has been solved by the use of carboxylate-containing aryl groups, resulting in the antitumour agent chlorambucil (Figure 27).
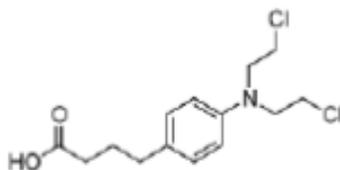
Figure 27: *Chlorambucil Structure of chlorambucil,*
a nitrogen mustard alkylating agent.

Ethyleneimines (aziridines)

Ethyleneimines are pre-formed aziridines and therefore constitute a natural extension of nitrogen mustards (Figure 28). To ensure antitumour activity, at least two ethyleneimine groups must be present in the molecule. To prevent protonation of the ethyleneimine, electron-withdrawing groups are attached (protonated ethyleneimines are too reactive). Lipophilic ethyleneimines are designed to enter the central nervous system.

Carboquone   Triethylenemelamine   Diaziquone
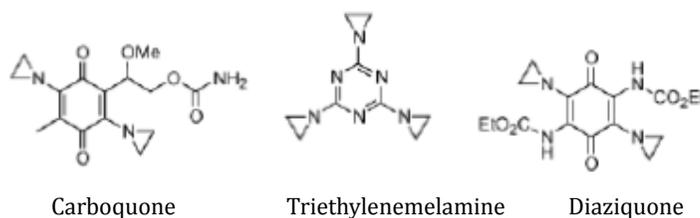
Figure 28: *Ethyleneimines Structure of ethyleneimines (aziridines),*
alkylating agents derived from nitrogen mustards.

Methanesulfonates

Methanesulfonates alkylate guanine at the N (7) position. Unlike nitrogen mustards, which tend to form inter-strand bridges, the methanesulfonates form intrastrand cross-links. An example is the bifunctional anti-cancer drug busulfan (Figure 29).
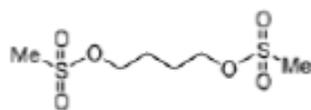
Figure 29: *Busulfan Structure of busulfan (or busulphan), an anti-cancer drug that forms intra-strand DNA cross-links.*

Platinum complexes

Cisplatin and carboplatin (Figure 30) represent a group of anti-cancer agents used in the treatment of testicular and ovarian tumours. Cisplatin and carboplatin form strong platinum-nitrogen bonds with guanine and adenine bases. The *cis* configuration leads to intra-strand cross-links, causing unwinding of the helix, preventing transcription and leading to cell death.
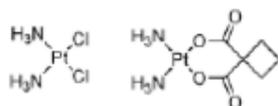
Figure 30: *Cisplatin and carboplatin Structures of cisplatin and carboplatin, alkylating agents used in the treatment of testicular and ovarian cancer.*

The *trans* isomer, *trans*-platin, is not an active anti-cancer agent, probably because it cannot readily form intra-strand cross-links. It tends to cross-link separate strands and such lesions are more easily repaired (Figure 31).
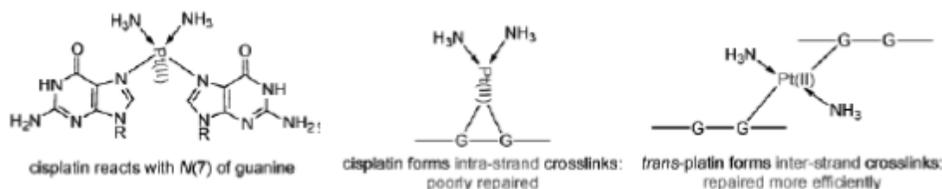
Figure 31: *The difference between cisplatin and trans-platin Cisplatin forms intra-strand cross-links, while trans-platin forms inter-strand cross-links, which are more readily repaired.*
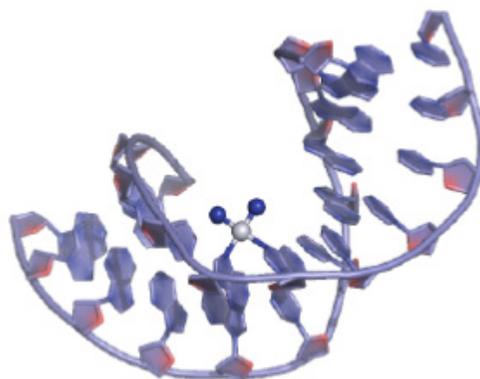
Figure 32: *Cisplatin DNA binding View from the three-dimensional structure of a cisplatin intra-strand adduct. The platinum atom is shown as a white sphere; the NH3 ligands are shown as blue spheres.*

DNA Cleavage Agents

Bleomycin

The classic DNA-cleaving anti-cancer antibiotic is bleomycin, a mixture of related antibiotics isolated from *Streptomyces verticillus*. Bleomycin is a glycopeptide antibiotic, the major component of which is bleomycin A_2 (Figure 33).
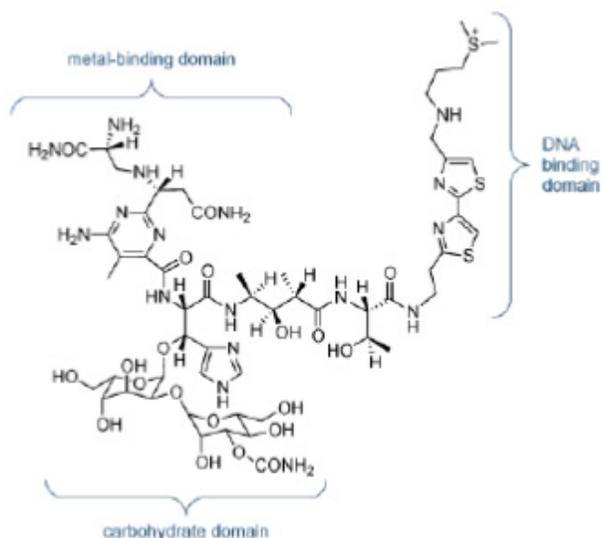
Figure 33: *Bleomycin Structure of Bleomycin A_2, an anti-cancer antibiotic.*

There are three domains in bleomycin A_2:

- Metal-binding domain: the pyrimidine, β-aminoalanine and β-hydroxyimidazole are involved in the formation of a stable complex with Fe (II). Reaction with O gives a ternary complex believed to be responsible for the DNA cleavage activity.
- DNA-binding domain: the bithiazole moiety intercalates into the double helix and the attached side chain containing a sulfonium ion is attracted to the phosphodiester backbone.
- Carbohydrate domain: the gluose and carbamoylated mannose disaccharide are thought to be responsible for selected accumulation of bleomycin in some cancer cells. This domain does not appear to be involved directly in DNA cleavage.

Bleomycin binds tightly to guanine bases in DNA, particularly in G-T and G-C-rich sequences. When the ternary complex of Fe (II), bleomycin and oxygen attacks DNA, it abstracts hydrogen atoms. The resultant radicals react with oxygen to form peroxy species which then fall apart, resulting in chain cleavage.

Enediyne antitumour antibiotics

The enediyne antitumour antibiotics are anti-cancer compounds that cause the oxygen-dependent cleavage of the DNA phosphate backbone. They interact with the minor groove of DNA and then either a thiol or NADPH triggers a reaction that produces radicals. These radicals cleave the DNA chain. The major characteristic of these enediynes is the presence of a macrocyclic ring with at least one double bond and two triple bonds.

The two steps in antitumour activity are therefore

- Activation of the enediyne
- Action of the activated antitumour agent on DNA.

Neocarzinostatin

The first of the enediyne antitumour agents, neocarzinostatin (Figure 34), was isolated in 1965. Further enediynes, principally esperamicins and calichemicins, were not isolated until many years later.
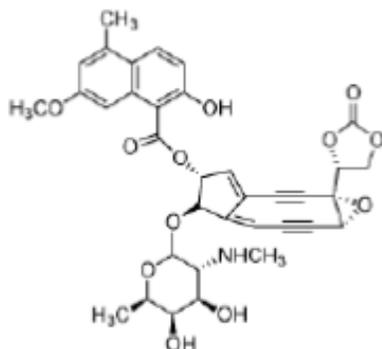
Figure 34: Neocarzinostatin Structure of neocarzinostatin, the first enediyne antitumour agent.

Neocarzinostatin contains an active naphthoate ester which intercalates in DNA, positioning the diyne in the minor groove. Activation with a thiol followed, by a Bergman rearrangement, produces a diradical that is somewhat different from those produced by other enediynes (it is not a 1, 4-dehydrobenzene diradical, but it is thought to behave similarly). The reactive diradical effects DNA strand scission by reaction with the C4'- and the C5'-atoms of the deoxyribose sugar, usually at deoxyadenosine and thymidine residues, consuming one equivalent of O_2 per strand break (Figure 35). As neocarzinostatin is the oldest member of the enediyne class of antitumour agents, it is the most thoroughly studied.
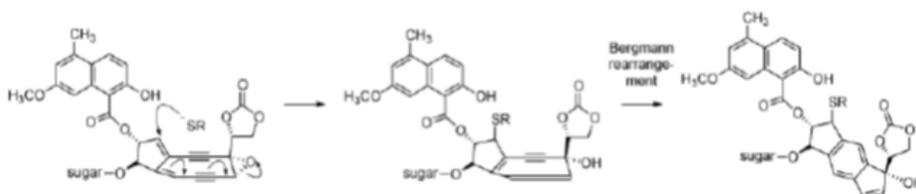
Figure 35: Neocarzinostatin mechanism Mechanism of activation of neocarzinostatin by thiols.

Esperamicins and calichemicins

The esperamicins and calichemicins are similar to neocarzinostatin, having a bicyclo [7.3.1] ring and an allylic trisulfide attached to the bridgehead carbon, a 3- ene-1, 5-diyne as part of the macrocycle and an α, β-unsaturated ketone in which the double bond is at the bridgehead of the bicyclic system (Figure 36).

Esperamicins   Calichemicins
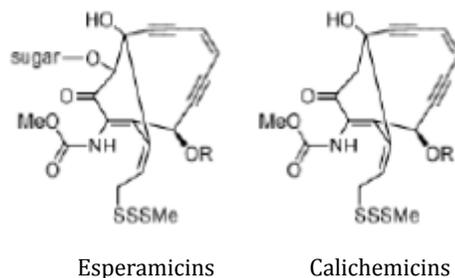
Figure 36: Esperamicins and calichemicins Structures of esperamicins and calichemicins, enediyne antitumour agents.

The proposed mechanism of activation of the esperamicins and calichemicins begins with reduction of the trisulfide to a thiolate by NADPH or thiols in the cell. This results in an intramolecular Michael addition, which followed by a Bergman rearrangement gives the 1, 4-dehydrobenzene diradical, the activated form of the esperamicin or calichemicin (Figure 37).
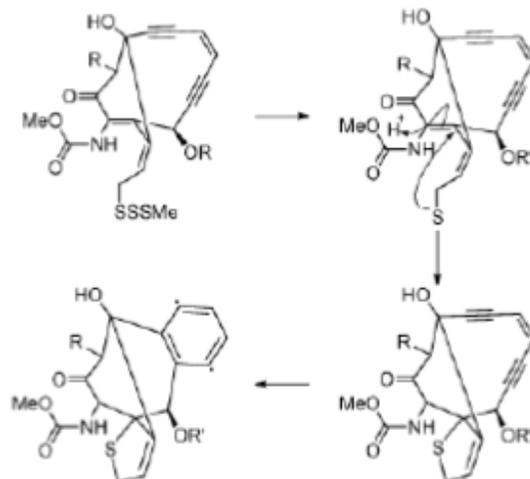
Figure 37: Esperamicin mechanism Mechanism of activation of esperamicins, enediyne antitumour agents.

Enediynes interact with the DNA minor groove, so that the pro-radical centers of the enediyne moiety are positioned close to the proton abstraction sites (Figure 38).
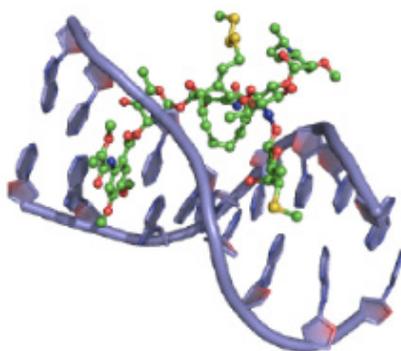
Figure 38: Esperamicin DNA binding View from the three-dimensional structure of a complex between esperamicin A1 (green) and a DNA duplex, showing the binding of esperamicin A1 in the minor groove of double-stranded DNA.

Dynemicin A

Dynemicin A (Figure 39) combines structural features of both anthracyclines and enediynes. Dynemicin A binds to DNA by a combination of intercalation and minor groove binding. Dynemicin A can be activated by NADPH, thiols or light, and, once bound to DNA, can cause cleavage of one or both strands of the DNA helix.
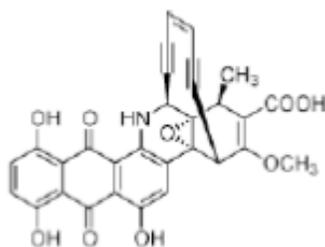
Figure 39: Dynemicin Structure of Dynemicin A, a small molecule that binds to DNA by intercalation and minor groove binding, and effects DNA strand cleavage.

The Structures of Quadruplex Nucleic Acids and Their Drug Complexes

Quadruplex nucleic acids may be defined as higher order structures formed by DNA or RNA sequences containing at least one contiguous tract of guanine nucleotides. The occurrence of such sequences in eukaryotic telomeres and within genomic sequences has been recognized for many years. More recent interest in the diverse structures formed by them has been catalyzed by the knowledge that they have highly distinctive features and thus may have utility as therapeutic targets [14-16] with inherently greater selectivity than duplex DNA.

This view has received support by the discovery and clinical development of a small molecule that has been found to interact with DNA quadruplexes within the ribosomal RNA template [17], as well as by the encouraging pre-clinical development of several compounds targeting telomeric DNA quadruplexes [15]. Since quadruplexes are three dimensional folded structures, the question has arisen as to whether the structures have sufficient diversity for exploitation by structure-based design methods, and indeed whether they are really druggable targets.

The underlying repeating motif of all quadruplexes is the G-quartet (also termed a G-tetrad), which has a high propensity to self-stack, with quadruplexes comprising a core of at least two stacked G-quartets [18, 19]. Quadruplexes have a requirement for metal ions [20], especially the alkali metals (and an order of preference $K^+ > Na^+$), and the ions are positioned in the interior channel that is formed at the centre of each G-quartet, coordinating to O6 atoms of the guanines in adjacent G-quartets.

It is this electrostatics feature that has made molecular dynamics simulations of DNA quadruplexes so much more challenging than those of duplex DNA [21]. The advent of newer force-field parameterisations for nucleic acids [22] coupled with use of the particle-mesh Ewald treatment of electrostatics has enabled significantly more stable simulations of quadruplexes to be achieved, but such calculations are still some way from being able to predict complete folding pathways for these structures [23].

Three principal categories of quadruplex arrangement are possible

1. Tetra molecular, with four separate strands associating together, with each strand having a minimum of one G-tract.
2. Bimolecular, with two strands, each normally having two G-tracts. A well-studied example is the sequence d (G4T4G4) from the telomeres of Oxytricha nova
3. Monomolecular (intramolecular), with normally four consecutive G-tracts. A number of natural sequences have been identified in which there are more than four G-tracts, and in those cases several distinct quadruplexes can be formed, often in equilibrium with each other.

The sequences intervening between successive G-tracts serve to link stacked G-quartets in a number of distinct ways, so that a variety of quadruplex topologies can be formed (Figure 40). These loops can be diagonal, lateral (also termed edgewise) or chain-reversal (also termed propeller), dependent on the number of G-quartets comprising a quadruplex, on loop length and sequence and sometimes on the nature of the alkali metal ion. Chain-reversal loops by definition connect two strands in the same parallel orientation, whereas diagonal and lateral loops connect chains in opposing, antiparallel orientations.

Native telomeric DNA and RNA quadruplexes

The terminal 100–200 nucleotides at the 30 end of human telomeres are normally single-stranded, so in principle telomeric DNA could readily fold into intramolecular DNA quadruplex structures. It is constrained from doing so by the hPOT1 single-stranded binding protein [24,25], and also in cancer cells, by the end-capping function of the human telomerase enzyme complex (comprising the catalytic domain hTERT + the RNA template recognition domain hTR). The concept of telomeric quadruplexDNA as a therapeutic target [14,15] was established with the finding that a group of disubstituted amidoanthraquinone small molecules, containing a planar aromatic chromophore, could inhibit telomerase activity [26].

The hypothesis underlying this observation was that a ligand molecule can induce the single-stranded telomericDNAsubstrate to fold into a quadruplex structure, which is known to be incompatible with telomerase-catalysed telomere elongation [27]. More recent studies have shown compelling evidence that the ligand effectively competes with hPOT1 and telomerase for the single-stranded overhang [28–30]. The formation of a quadruplex–ligand complex at telomere ends appears to be equivalent to the exposure of damaged DNA, since it elicits a rapid DNA damage response that is lethal to the affected cells [31,32].

The design of quadruplex-binding ligands has not to date, with few exceptions, been based on detailed experimental structural data. Partly, this is a consequence of the lack of such data until very recently. In addition, structural studies on native human telomeric quadruplexes ie those containing two or more repeats of d(TTAGGG), which could be a starting-point for structure-based design, have revealed that these structures can be highly polymorphic, in accord with a number of biophysical studies in solution [33,34]. Therefore it has not been obvious as to which of the determined structures could be used as a template for drug design.
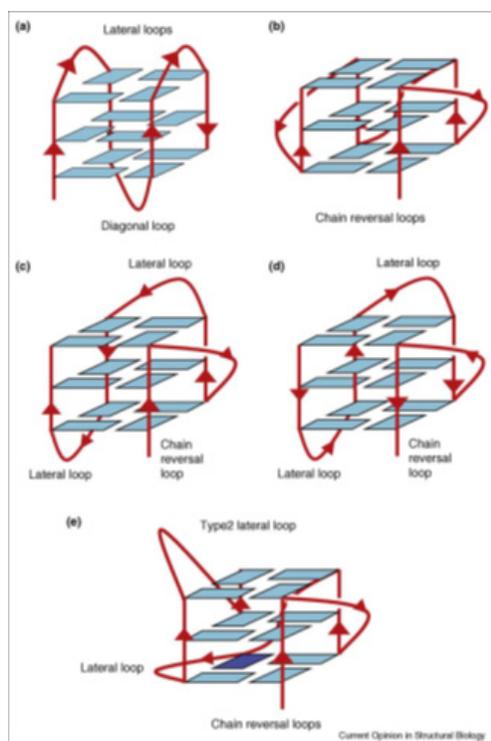
Figure 40: *Schematic views of various intramolecular quadruplex topologies. (a) Antiparallel topology as found in the Na+ form of a human telomericquadruplex [35]. (b) Parallel fold as found in the crystal structure of a K+ form of a human telomeric quadruplex [36]. (c) (3 + 1) hybrid topology (form 1) as found in the NMR structures of the K+ form of human telomeric quadruplex sequences [37, 38]. (d) (3 + 1) hybrid topology (form 2) as found in the NMR structures of the K+ form of human telomeric quadruplex sequences [39,40]. Note that the order of loops is now reversed compared with that in form 1. (e) Topology of the c-kit promoter quadruplex [41].*

To date almost all studies on telomeric quadruplexeshave focussed on those formed from DNA sequences. This emphasis is starting to change with the unexpected finding that telomeric DNA can be transcribed into RNA that is then fragmented into shorter telomeric RNA sequences that appear to have telomerase regulatory activity [42,43]. No proteins that bind to these RNAs have as yet been reported. A comprehensive biophysical analysis of the single-repeat and two-repeat sequences r [UUAG$_3$] and r [UAG3UUAG$_3$U] has shown [44,45] that these form unique tetramolecular and bimolecular quadruplex respectively, with all strands parallel and having chain-reversal loops. Remarkably this topology is formed both in Na$^+$ and K$^+$ solution.

Quadruplex ligand complexes

A large number of ligands have been examined subsequent to the original finding of quadruplex-mediated telomerase inhibition by amidoanthraquinone derivatives [26], and elaborated as telomeric quadruplex ligands, mostly with a view to their eventual development as anticancer agents [15,16,46]. The underlying features of their structure-activity relationships has been more elusive, beyond the basic requirements shown by most such ligands, of (i) a planar aromatic chromophore, and (ii) substituents with terminal basic groups.

Early NMR and X-ray studies on structurally simple tetramolecular quadruplex ligand complexes (such as in the crystal structure formed by four strands of d (TGGGGT) associating with the drug daunomycin [47], concurred in establishing that the chromophores stack onto the terminal G-quartets of these quadruplexes, and thus that binding between internal G-quartets analogous to intercalation between successive base pairs, as observed in duplex DNA, does not occur. The G-quartet end-stacking binding mode was also observed in the crystal structure of a 1:1 complex between a disubstituted acridine ligand and the dimeric antiparallel Oxytricha nova telomeric quadruplex (Figure 41).

A TMPyP4 molecule is sandwiched between two such base pairs that are stacked between individual quadruplexes in the crystal (Figure 42). Three more recent crystal structures of telomeric quadruplex– ligand complexes have similarly revealed parallel topologies and end-binding modes, but with more predictable and direct ligand-G-quartet stacking. Thus, the structure [48] of a complex with the experimental drug BRACO-19 shows a drug molecule sandwiched between two bimolecular quadruplexes to form the biological unit (Figure 41).
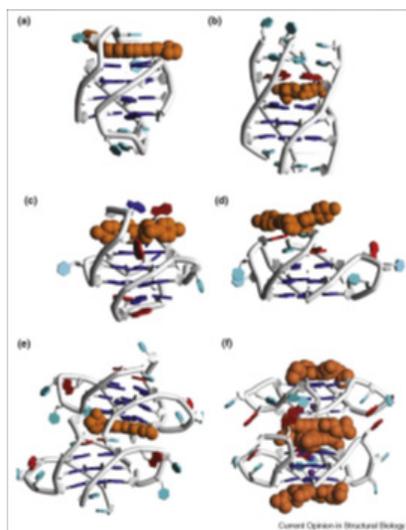
Figure 41: Cartoon representations of quadruplex–ligand complex structures, drawn from coordinates deposited in the PDB. PDB id codes. The ligand is shown in space-filling representation in each structure, colored orange. (a) Structure 1L1H [49]. Ligand is a disubstituted amidoacridine molecule. (b) Structure 1NMZ [50]. Ligand is the experimental drug RHPS4. (c) Structure 2A5R [51]. Ligand is the porphyrin molecule TMPyP4. (d) Structure 2HR1 [52]. Ligand is the porphyrin molecule TMPyP4. (e) Structure 3CE5 [48]. Ligand is the experimental drug BRACO-19. (f) Structure 3CDM [53]. Ligand is a naphthalene diimide derivative.
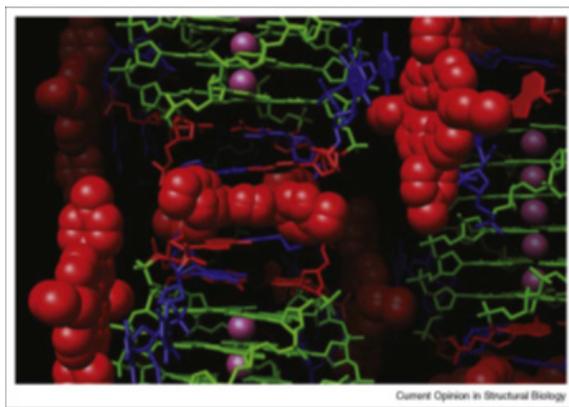
Figure 42: *A view of the crystal structure of the bimolecular telomeric quadruplex with bound TMPyP4 (shown in red space-filling mode). A TMPyP4 molecule is shown intercalated between reverse Watson-Crick base pairs (in red and blue stick form respectively), within a column of quadruplexes.*

Genomic quadruplexes

Application of informatics methods to human [54,55] and other genome sequences [56–59] has been used to quantitate and extend earlier findings that potential quadruplex sequences are not restricted to telomeres and can occur at numerous genomic loci, including those in yeast and bacterial genomes. There is a significant bias towards occurrence in promoter regions [57-61] and 50 UTRs [62]. The searches have used the definition of an intramolecular quadruplex-forming sequence as: $G_{3-5}N_{L1}G_{3-5}N_{L2}G_{3-5}N_{L3}G_{3-5}$ where N_{L1-3} are loops of unknown length, but within the (arbitrary) limits of $1 < N_{L1-3} < 7$ nucleotides.

The two analyses of quadruplex sequence prevalence in the human genome, which used distinct algorithms [54,55], have arrived at closely similar conclusions, on the basis of the above sequence definition, that there are ca 300–400 000 quadruplex sequence occurrences ('putative quadruplexes'). It is currently not possible to predict the topology or relative stability of any given sequence, and the majority of studies to date have focused on a few examples, with such sequences being over-represented in oncogene promoter regions such as that of the c-myc oncogene. Analyses, both experimental and theoretical, have suggested on the basis of frequency of occurrence [54] that short loop sequences are themselves significantly over represented and that the occurrence of ≥ 2 such single-nucleotide loops results in a parallel topology [63–65]. Much of the current interest in genomic quadruplexes is based on the concept that small molecules with quadruplex selectivity would stabilize quadruplex formation in, for example, a promoter sequence, which would then downregulate gene expression [66,67].

Solvent-Accessible Surfaces of Proteins and Nucleic Acids

Computer graphics has made the results of x-ray crystallographic studies of proteins and nucleic acids more accessible to biochemists and molecular biologists. Traditionally, computer-generated images of molecular structures have consisted of lines for the chemical bonds [68-70] or spheres [71-74] and ellipsoids [75] for the atoms. A method is presented for analytically calculating a smooth, three-dimensional contour about a molecule. The molecular surface envelope may be drawn on either color raster computer displays or real-time vector computer graphics systems. Molecular areas and volumes may be computed analytically from this surface representation. Unlike most previous computer graphics representations of molecules, which imitate wire models or space-filling plastic spheres, this surface shows only the atoms that are accessible to solvent.

This analytical method extends the earlier dot surface numerical algorithm, which has been applied in enzymology, rational drug design, immunology, and understanding DNA base sequence recognition. Applications of this surface representation include enzymology, rational drug design, and the elucidation of molecular dis-eases such as sickle cell anemia, recognition of specific DNA base sequences by proteins and drugs, and the location of possible antigenic determinants on viruses. The historical basis for the smooth

surface envelope method is the work of Richards [76] and colleagues on solvent-accessible area. Their emphasis was on chemical calculations measuring quantities of hydrophobic and hydrophilic area, while the methods described below were developed primarily for the purpose of visualizing molecular structure and interactions. Nevertheless, these new methods also permit the measurement of area and volume in conjunction with the graphical display.

Solvent-Accessible Area

Solvent-accessible area was originally defined and computed by Lee and Richards [77] as the area traced out by the center of a probe sphere representing a solvent molecule as it is rolled over the surface of the molecule of interest. These computational methods were invented as a tool for attacking the protein folding problem [76]. The problem is that of predicting the three-dimensional structure of a protein given only its primary sequence of amino acids. Simply measuring a quantity of area is insufficient for the study of many aspects of protein and nucleic acid function, such as substrate binding and catalysis, drug-nucleic acid interaction, and recognition by the immune system. A method for visualizing solvent-accessible surreentrant surfaces and of Shrake and Rupley [78] for calculating solvent-accessible area, I developed a numerical computer algorithm for placing dots over the solvent-accessible molecular surface of a protein [79,80]. Below, I briefly review the dot surface algorithm and present some, analytical surface methods.

For each probe position that does not experience van der Waals over-lap with the atoms of the protein, points lying on the inward-facing surface of the probe sphere become part of the pro-tein's solvent-accessible surface. The probe may be placed tangent to (i) single atoms, creating a dot at the point of tangency, (ii) pairs of atoms, creating a concave arc of dots connecting the two points of tangency, and (iii) triples of atoms, creating a concave triangle of dots between the three points of tangency. For each surface point generated, the numerical algorithm produces not only its coordinates but also an approximate solvent-accessible area associated with the point and an outward-pointing unit vector perpendicular to the surface at that point. The pancreatic trypsin-tryp-sin inhibitor complex [81] is shown in (Figure 43), with a dot surface for the enzyme only.
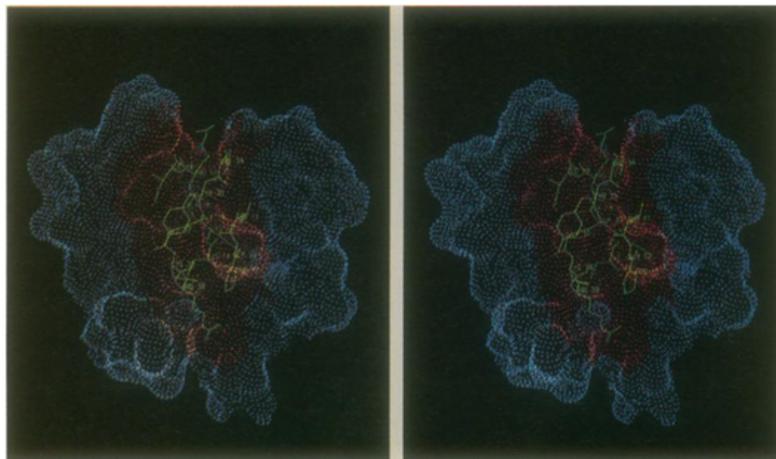
Figure 43: Stereo pair of the pancreatic trypsin-trypsin inhibitor complex. The enzyme is represented by a dot surface. The residues of the inhibitor in contact with the enzyme are represented by bonds. The part of the trypsin surface that is kept from contact with the solvent by the presence of the inhibitor is colored red.

This method has proved useful in enzymology [82-86], immunology [87,88], virology [89], molecular pathology [90], and the study of protein-ligand [91] and protein-protein [92-94] interactions. Despite the many applications of the dot surface numerical algorithm, it was necessary to invent an analytical surface algorithm in order to generate high-resolution color raster display images and to compute more accurate molecular areas and volumes. A continuous molecular surface contour is defined as the union of pieces of spheres and tori joining smoothly at circular arcs. There are three kinds of pieces: concave spherical triangles, saddle-shaped rectangles, and convex spherical regions (Figure 44).
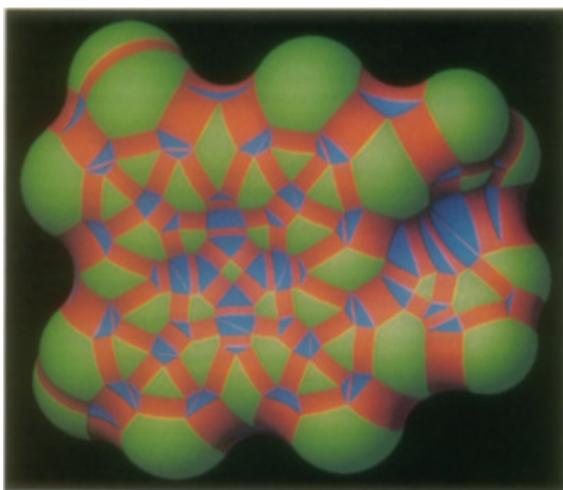
Figure 44: *Heme molecule drawn on a color raster graphics system. Green, convex surface;*
red, saddle surface; blue, concave surface. Surface pieces join at circular arcs.

Each concave triangle has three concave arcs as edges. Next, the saddle rectangles are formed by con-necting adjacent concave arcs along the inner surfaces of tori (Figures 45 and 46). The edges of each saddle rectangle consist of a pair of concave arcs and a pair of convex arcs. In the final step, the convex arcs on each atom are grouped to form closed circuits, or cycles, and the boundary of each convex face is defined by zero, one, or more cycles.
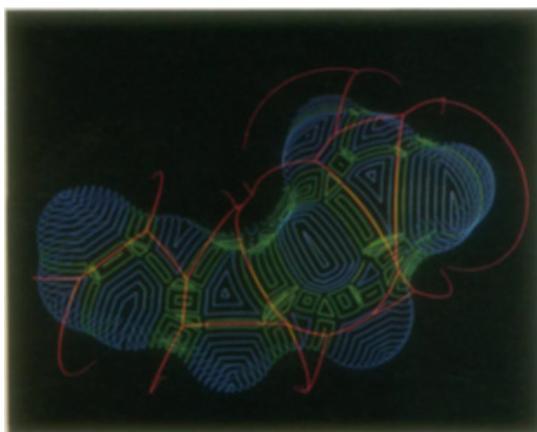
Figure 45: *Trajectory of probe rolling over a molecular surface. The trajectory arcs (red) con-*
nect positions where the probe is simultaneously tangentt o three atoms. In a correspondingm
anner,s addler ectanglesc onnect concave triangles. T hese reentrants urfaces (green)t hen
define the boundaries of the convex surfaces (magenta).

Nucleic Acid Medicine

"Nucleic acid medicines" utilize "nucleic acid," which refer to substances such as DNA and RNA that control genetic information, as drugs. These allow targeting of molecules such as mRNA and mi RNA that cannot be targeted with traditional low molecular weight drugs and antibody medicines, and there is a great expectation for these drugs as next generation pharmaceuticals. Active research is being conducted globally as it is expected to lead to the creation of drugs which were previously intractable.
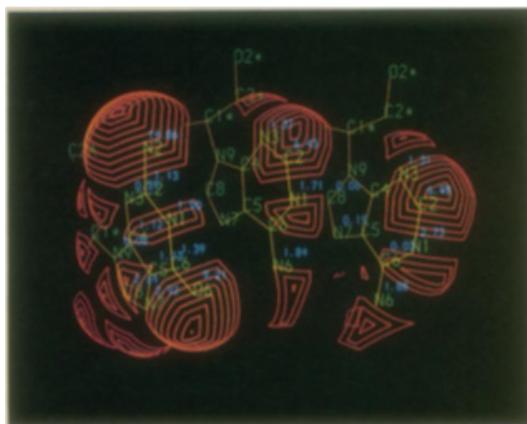
Figure 46: *Y east phenylalanylt ransferR NA anticodon (GAA). The contact surface of the three anticodon bases is shown. The contact areas in square angstroms are displayed next to the atom labels.*

On the other hand, it has been pointed out that the development of nucleic acid medicines has issues to overcome, including "(i) instability of nucleic acid molecules in the body," "(ii) concerns for adverse drug reactions," and "(iii) difficulty in the drug delivery system (DDS)." Also, Japanese companies are steps behind in the development of nucleic acid medicines due to the monopolization of dominant patents of nucleic acid by companies in Europe and the US, causing interference with Japanese development.

Characteristics of nucleic acid medicines

Nucleic acid medicines are next generation drugs [95], which use "nucleic acids" such as DNA and RNA. While research are being conducted on nucleic acid medicines in various countries around the world, Bonac has developed "Bonac Nucleic Acid" based on a novel approach. New nucleic acid medicines developed by Bonac.

Bonac Nucleic Acid

Bonac Nucleic Acid [96] is characterized by its unique structure of single-stranded long-chain nucleic acid molecules, which is different than the previously used siRNA. This structure not only improved stability of the drug in the body, but also enabled avoidance of innate immune response mediated by Toll-like receptor 3 (TLR3). Having overcome the issues associated with the previously used siRNA, Bonac Nucleic Acid is expected to become a new type of nucleic acid medicine.

What is Bonac Nucleic Acid?

Definition of Bonac Nucleic Acid

"Bonac Nucleic Acid" is a generic name for a new single-stranded nucleic acid discovered by Bonac. This includes the substances covered by dominant patient as having a specific structure, and this specific structure is in charge of the unique function of Bonac Nucleic Acid.

Nucleic acid is a component essential to the body. However, since nucleic acid must be digested and metabolized for use in the body, it is immediately degraded in the body and therefore is unstable. Bonac Nucleic Acid is made from a natural form of nucleic acid; however, it is resistant to degradation due to its characteristic structure and can be delivered in a stable manner to the area of the body that requires the drug. In addition, since Bonac Nucleic Acid is made from a natural form of nucleic acid, it is considered safe to the body. Due to the design of the internal sequence in Bonac Nucleic Acid, a similar effect to siRNA and miRNA can be obtained. We are also continuing to meet challenges in new areas where the technology of Bonac Nucleic Acid can be utilized.

Classification of Bonac Nucleic Acid

Bonac Nucleic Acid has different shapes and functions depending on the functionality given.
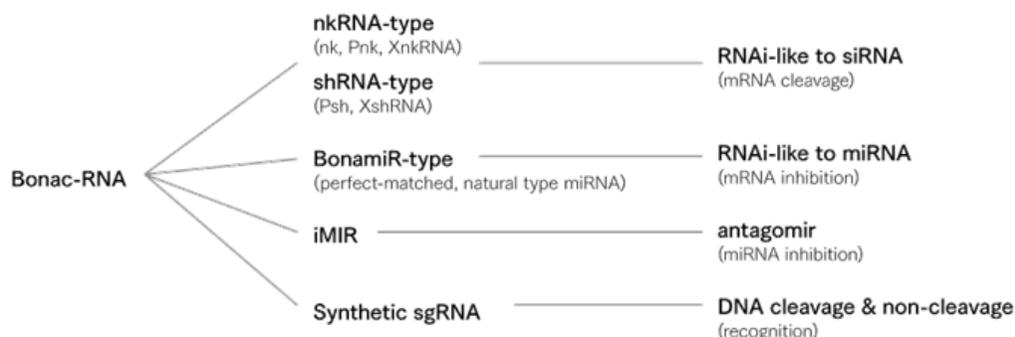
Structure of Bonac Nucleic Acid

- Single-stranded nucleic acid with chain length not less than 50 bases
- Unique secondary structure due to self-annealing, and annealing process is not required
- Devised not only nucleic molecules but also unique hybrid molecule (PnkRNA) with unique amino acid amidite (ingredient for nucleic acid)
- Higher stability compared to traditional method
- Does not trigger immune response

Traditional RNA

- Double-stranded nucleic acid with each chain consisting of not more than 49 bases
- Requires annealing process
- Susceptible to degradation and measures must be taken to stabilize
- May trigger immune response

"Nucleic acid medicine" is a next generation drug discovery technology [95] with a completely different mechanism of action than traditional pharmaceutical products. It also features the ability to be manufactured easily at moderate sized molecules and the potential to exhibit efficacy and safety that surpasses those of antibody medicines. Due to these features, there is an expectation for nucleic acid medicines to be applied in cancer and hereditary disorders which were previously difficult to treat, as well as in illnesses such as influenza and viral infections.
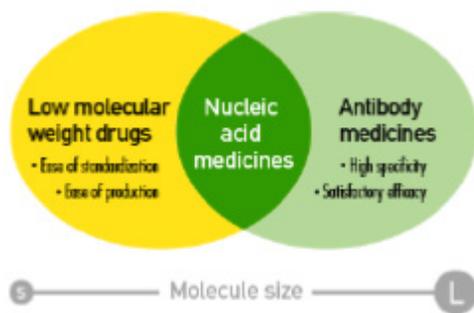
Types of nucleic acid medicines

Nucleic acid medicines which utilizes DNA and RNA include those that target nucleic acids at the stage where protein is synthesized from genome DNA (such as mRNA and miRNA) and those that target protein.
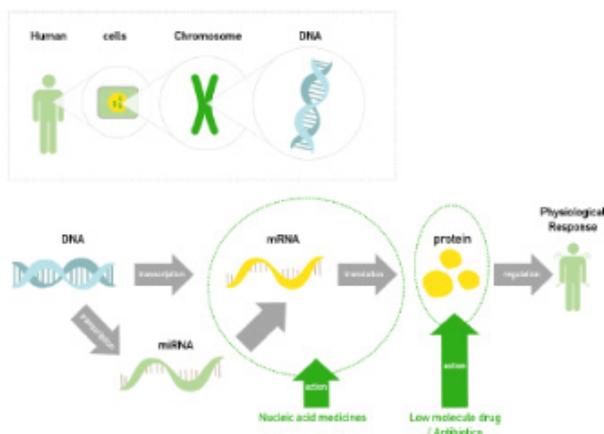
Types and characteristics of nucleic acid medicines (drugs for prophylaxis and treatment)

There are nucleic acid medicines with different types and characteristics according to the targets and mechanisms of action.

Type	Target	Site of action	Mechanism of action	Summary
siRNA	mRNA	Inside the cell (cytoplasm)	mRNA cleavage	Double-stranded RNA with cleavage of mRNA homologous to the sequence (siRNA), single stranded hairpin RNA (shRNA), etc. with effect according to the principle of RNAi
miRNA	microRNA	Inside the cell (cytoplasm)	microRNA replacement	Double-stranded RNA, miRNA of single-stranded hairpin RNA or its mimic is used to strengthen the function of miRNA deteriorated by disorders
Antisense	mRNA, miRNA	Inside the cell (in the nucleus cytoplasm)	mRNA and miRNA degradation, splicing inhibition	Single-stranded RNA/DNA which binds to the target mRNA and miRNA to cause degradation or inhibition, or acts to skip exon when splicing

Aptamer	Protein (extracellular protein)	Outside the cell	Functional inhibition	Single-stranded RNA/DNA which binds to the target protein in a similar manner to antibodies/DNA
Decoy	Protein (transcription factor)	Inside the cell (in the nucleus)	Transcription inhibition	Double-stranded DNA with identical sequence to the binding site for transcription factor, which binds to the transcription factor of the affected gene to suppress the target gene
Ribozyme	RNA	Inside the cell (cytoplasm)	RNA cleavage	Single-stranded RNA with enzyme function for binding and cleavage of target RNA
CpG oligo	Protein (receptor)	Cell surface	Immunopotentiation	Oligodeoxynucleotide with CpG motif (single-stranded DNA)
Other	____	_____	_____	Nucleic acid medicines other than those listed above which act to activate innate immunity, such as PolyI : PolyC (double-stranded RNA) and antigen

Table 1: *Types and characteristics of nucleic acid medicines.*

Mechanism of siRNA silencing
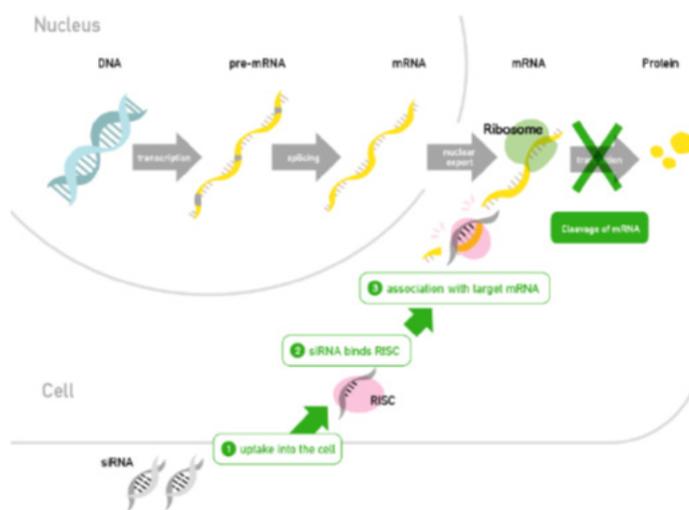
Nucleic Acids as Therapeutics

The thickening of the vessel wall (intimal hyperplasia) is a pathological process [97] which often follows revascularization approaches such as transluminal angioplasty and artery bypass graft, procedures used to re-vascularize stenotic artery. Despite the significant improvements in the treatment of intimal hyperplasia obtained in the last years, the problem has not completely solved.

Nucleic acid based-drugs (NABDs) represent an emergent class of molecules with potential therapeutic value for the treatment of intimal hyperplasia. NABDs of interest in the field of intimal hyperplasia are: ribozymes, DNAzymes, antisense oligonucleotides, decoy oligonucleotides, small interfering RNAs and micro interfering RNAs. These molecules can recognize, in a sequence specific fashion, a target which, depending on the different NABDs, can be represented by a nucleic acid or a protein. Upon binding, NABDs can down-modulate the functions of the target (mRNA/proteins) and thus they are used to impair the functions of disease-causing biological

molecules. In spite of the great therapeutic potential demonstrated by NABDs in many experimental model of intima hyperplasia, their practical use is hindered by the necessity to identify optimal delivery systems to the vasculature.

What is a gene?

A gene is the basic physical and functional unit [7] of heredity. Genes, which are made up of DNA, act as instructions to make molecules called proteins. In humans, genes vary in size from a few hundred DNA bases to more than 2 million bases. The Human Genome Project has estimated that humans have between 20,000 and 25,000 genes. Every person has two copies of each gene, one inherited from each parent. Most genes are the same in all people, but a small number of genes (less than 1 percent of the total) are slightly different between people. Alleles are forms of the same gene with small differences in their sequence of DNA bases. These small differences contribute to each person's unique physical features.
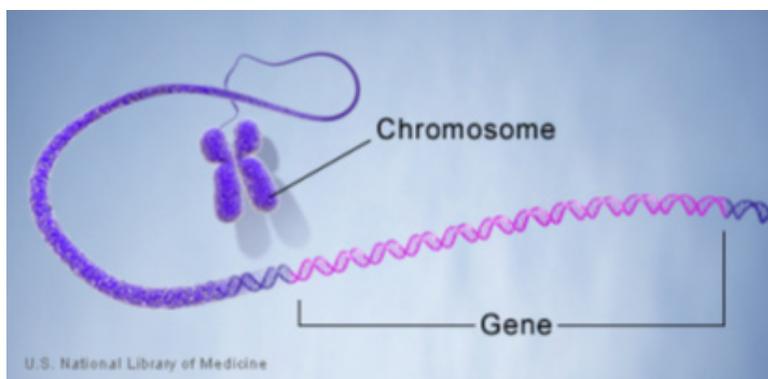
Figure 47: *Genes are made up of DNA. Each chromosome contains many genes.*

For more information about genes

Genetics Home Reference provides consumer-friendly gene summaries (https://ghr.nlm.nih.gov/gene) that include an explanation of each gene's normal function and how mutations in the gene cause particular genetic conditions. The Centre for Genetics Education offers a fact sheet that introduces genes and chromosomes (http://www.genetics.edu.au/publications-and-resources/factssheets/fact-sheet-1-an-introduction-to-dna-genes-and-chromosomes). The Tech Museum of Innovation at Stanford University describes genes and how they were discovered (http://genetics.thetech.org/about-genetics/what-gene). The Virtual Genetics Education Centre, created by the University of Leicester, offers additional informationonDNA, genes, and chromosomes (http://www2.le.ac.uk/projects/vgec/schoolscolleges/topics/dna-genes-chromosomes).

What is a chromosome?

In the nucleus of each cell, the DNA molecule is packaged into thread-like structures called chromosomes. Each chromosome is made up of DNA tightly coiled many times around proteins called histones that support its structure. Chromosomes are not visible in the cell's nucleus-not even under a microscope-when the cell is not dividing. However, the DNA that makes up chromosomes becomes more tightly packed during cell division and is then visible under a microscope. Most of what researchers know about chromosomes was learned by observing chromosomes during cell division. Each chromosome has a constriction point called the centromere, which divides the chromosome into two sections, or "arms." The short arm of the chromosome is labeled the "p arm." The long arm of the chromosome is labeled the "q arm." The location of the centromere on each chromosome gives the chromosome its characteristic shape, and can be used to help describe the location of specific genes.
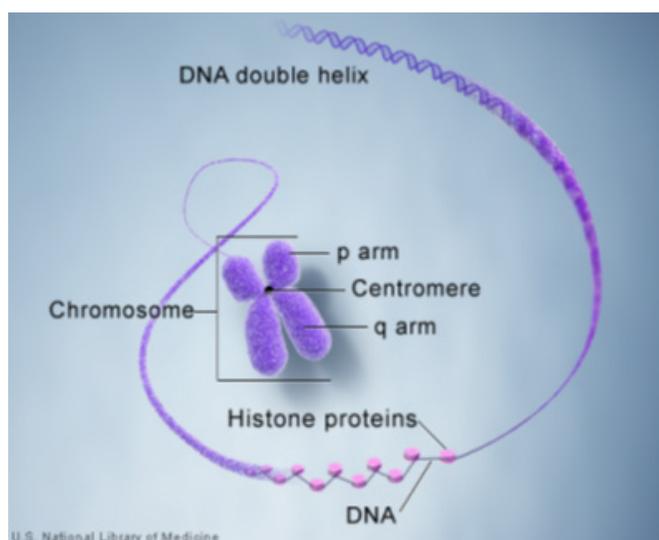
Figure 48: DNA and histone proteins are packaged into structures called chromosomes.

For more information about chromosomes

Genetics Home Reference provides information about each human chromosome (https://ghr.nlm.nih.gov/chromosome) written in lay language. The Centre for Genetics Education offers a fact sheet that introduces genes and chromosomes (http://www.genetics.edu.au/publications-and-resources/factssheets/fact-sheet-1-an-introduction-to-dna-genes-and-chromosomes). GeneEd also providesinformationaboutthebasicsofchromosomes https://geneed.nlm.nih.gov/topic_subtopic.php?tid=15&sid=17. The University of Utah's Genetic Science LearningCenteroffersadescriptionofchromosomes (http://learn.genetics.utah.edu/content/basics/readchromosomes/), including how scientists tell them apart.

What is gene therapy?

Gene therapy [7,98] is an experimental technique that uses genes to treat or prevent disease. In the future, this technique may allow doctors to treat a disorder by inserting a gene into a patient's cells instead of using drugs or surgery. Researchers are testing several approaches to gene therapy, including:

• Replacing a mutated gene that causes disease with a healthy copy of the gene.
• Inactivating, or "knocking out," a mutated gene that is functioning improperly.
• Introducing a new gene into the body to help fight a disease.

Although gene therapy is a promising treatment option for a number of diseases (including inherited disorders, some types of cancer, and certain viral infections), the technique remains risky and is still under study to make sure that it will be safe and effective. Gene therapy is currently only being tested for the treatment of diseases that have no other cures.

Thus, gene therapy is a technique [98] for correcting defective genes responsible for disease development. Nucleic acid-based molecules (deoxyribonucleic acid, complementary deoxyribonucleic acid, complete genes, ribonucleic acid, and oligonucleotides) are utilized as research tools within the broad borders of gene therapy and the emerging field of molecular medicine. Although most of the nucleic acid-based drugs are in early stages of clinical trials, these classes of compounds have emerged in recent years to yield extremely promising candidates for drug therapy to a wide range of diseases, including cancer, infectious diseases, diabetes, cardiovascular, inflammatory, and neurodegenerative diseases, cystic fibrosis, hemophilia, and other genetic disorders. Gene therapy may be classified into two types: somatic and germ line gene therapy.

Genetic mutation and cancer development
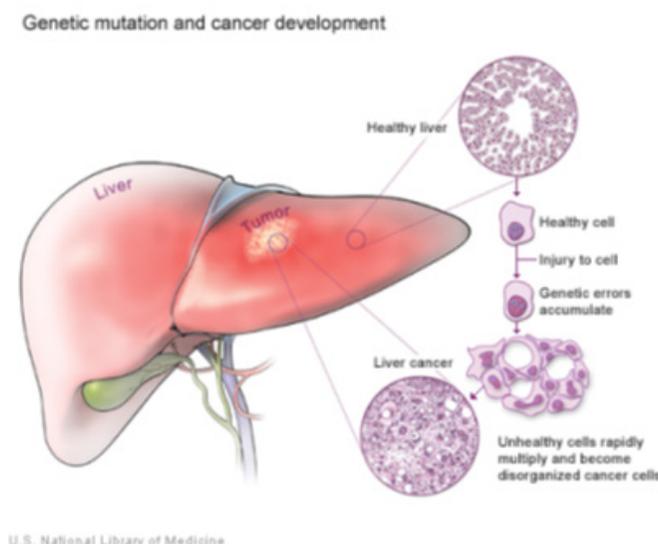
U.S. National Library of Medicine

Figure 49: Cancer results when cells accumulate genetic errors and multiply without control.

There are many ethical, social, and commercial issues raised by the prospects of treating patients using gene therapy. This part summarizes deoxyribonucleic acid-based therapeutics, ribonucleic acid-based therapeutics, and gene transfer technologies. Deoxyribonucleic acidbased therapeutics includes plasmids, oligonucleotides for antisense and antigene applications, deoxyribonucleic acid aptamers, and deoxyribonucleic acidzymes, while ribonucleic acid-based therapeutics includes ribonucleic acid aptamers, ribonucleic acid decoys, antisense ribonucleic acid, ribozymes, small interfering ribonucleic acid, and micro ribonucleic acid. This part also includes current status of gene therapy and recent developments in gene therapy research.

Modern drug research aims to discover biologically active molecule(s) that are absolutely specific to the molecular targets responsible for the disease progression. Moreover, there is strong belief that medicine will soon benefit from the development of new therapeutic technologies to directly target human genes. Insertion of new genetic material into the cells of an individual with the intention of producing a therapeutic benefit for the patient is human gene therapy [99], while gene therapy is a technique for correcting defective genes responsible for disease development. Numerous gene therapy strategies are under development, some of which use nucleic acid-based molecules to inhibit gene expression at either the transcriptional or posttranscriptional level, and this strategy has potential applications, such as in cardiovascular [100]; Patil and Burgess [101] and inflammatory disorders, cancer [102], neurological disorders [103], and infectious diseases, as well as in organ transplantation.

A number of human diseases are known to be genetic in origin (e.g., Huntington's chorea and cystic fibrosis), and virtually all diseases, except for some trauma, have a hereditary component [104]. Thus, gene therapy represents an opportunity for the treatment of genetic disorders in humans by modifying their cells genetically [105].

Despite the obvious advantages [98], that might be gained from human gene therapy (i.e., replacing a defective gene with a normal one), there are many ethical, social, and commercial issues surrounding the technology. The outcome of an error in technology might not be observed for many years. Moreover, it is feared that unpredictable and perhaps irreversible side effects occur in treated individuals. The social implications of such technology include the possibility that patients might suffer from depression as a result of being "genetically altered" or might not be accepted by society in the way that they were before treatment. The commercial implications of such technology are that the insurance companies and other such institutions also would want to access the available information prior to them granting life insurance policies, etc. Hence, it is obvious that a person shown to have a predisposition to a genetic disease could be severely penalized because of a mutation in their DNA.

The possibility of using nucleic acids as drugs for the treatment of genetic diseases is still very much in its infancy. One reason for this is that, unlike monogenetic disorders such as severe combined immune deficiency (SCID), which is caused by a mutation in the adenosine deaminase (ADA) gene, very few diseases are caused by a single gene mutation; most are caused by the mutation of multiple genetic components. For example, cancer usually involves multiple genetic lesions within the same cell and it is unlikely that the nature of every one of these oncogenic mutations is yet known.

Elucidation of the human genome has also provided a major impetus in identifying human genes implicated in diseases, which may eventually lead to the development of nucleic acid-based drugs for gene replacement or potential targets for gene ablation [106]. Moreover, the Human Genome project will help determine genetic markers responsible for patient response to drug therapy, drug interactions, and potential side effects [107]. Currently, all gene therapy trials approved for human use target somatic cells that will live only as long as the patient, and this is known as the somatic gene therapy.

Its purpose is to alleviate disease in the treated individual alone. In contrast, it is also possible to target directly the gametes (sperm and ova) to modify the genetic profile of the subsequent generation of unborn "patients." This gene transfer at an early stage of embryonic development is known as the germ line gene therapy. More than 300 clinical trials involving gene transfer in patients have been approved, and the first nucleic acid drug, an antisense oligonucleotide, fomivirsen (marketed as Vitravene), has been approved by the US Food and Drug Administration (FDA) for the treatment of cytomegalovirus retinitis in immunocompromised patients [108].

How does gene therapy work?

Gene therapy is designed to introduce genetic material [7] into cells to compensate for abnormal genes or to make a beneficial protein. If a mutated gene causes a necessary protein to be faulty or missing, gene therapy may be able to introduce a normal copy of the gene to restore the function of the protein. A gene that is inserted directly into a cell usually does not function. Instead, a carrier called a vector is genetically engineered to deliver the gene.

Certain viruses are often used as vectors because they can deliver the new gene by infecting the cell. The viruses are modified so they can't cause disease when used in people. Some types of virus, such as retroviruses, integrate their genetic material (including the new gene) into a chromosome in the human cell. Other viruses, such as adenoviruses, introduce their DNA into the nucleus of the cell, but the DNA is not integrated into a chromosome.

The vector can be injected or given intravenously (by IV) directly into a specific tissue in the body, where it is taken up by individual cells. Alternately, a sample of the patient's cells can be removed and exposed to the vector in a laboratory setting. The cells containing the vector are then returned to the patient. If the treatment is successful, the new gene delivered by the vector will make a functioning protein.

Researchers must overcome many technical challenges before gene therapy will be a practical approach to treating disease. For example, scientists must find better ways to deliver genes and target them to particular cells. They must also ensure that new genes are precisely controlled by the body. A new gene is injected into an adenovirus vector, which is used to introduce the modified DNA into a human cell. If the treatment is successful, the new gene will make a functional protein.

Is gene therapy safe?

Gene therapy is under study to determine whether it could be used to treat disease. Current research is evaluating the safety of gene therapy; future studies will test whether it is an effective treatment option. Several studies have already shown that this approach can have very serious health risks, such as toxicity, inflammation, and cancer. Because the techniques are relatively new, some of the risks may be unpredictable; however, medical researchers, institutions, and regulatory agencies are working to ensure that gene therapy research is as safe as possible.
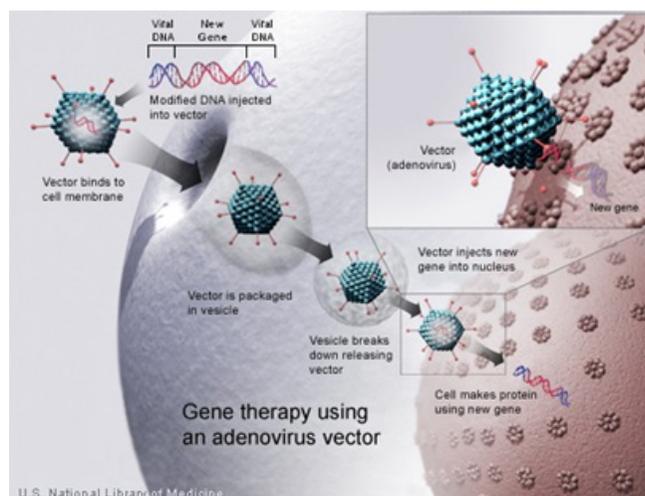
U.S. National Library of Medicine

Figure 50: *A new gene is injected into an adenovirus vector, which is used to introduce the modified DNA into a human cell. If the treatment is successful, the new gene will make a functional protein. Credit: U.S. National Library of Medicine.*

Comprehensive federal laws, regulations, and guidelines help protect people who participate in research studies (called clinical trials). The U.S. Food and Drug Administration (FDA) regulates all gene therapy products in the United States and overseas research in this area. Researchers who wish to test an approach in a clinical trial must first obtain permission from the FDA. The FDA has the authority to reject or suspend clinical trials that are suspected of being unsafe for participants.

The National Institutes of Health (NIH) also plays an important role in ensuring the safety of gene therapy research. NIH provides guidelines for investigators and institutions (such as universities and hospitals) to follow when conducting clinical trials with gene therapy. These guidelines state that clinical trials at institutions receiving NIH funding for this type of research must be registered with the NIH Office of Biotechnology Activities.

The protocol, or plan, for each clinical trial is then reviewed by the NIH Recombinant DNA Advisory Committee (RAC) to determine whether it raises medical, ethical, or safety issues that warrant further discussion at one of the RAC's public meetings.

An Institutional Review Board (IRB) and an Institutional Biosafety Committee (IBC) must approve each gene therapy clinical trial before it can be carried out. An IRB is a committee of scientific and medical advisors and consumers that reviews all research within an institution. An IBC is a group that reviews and approves an institution's potentially hazardous research studies. Multiple levels of evaluation and oversight ensure that safety concerns are a top priority in the planning and carrying out of gene therapy research.

What are the ethical issues surrounding gene therapy?

Because gene therapy involves making changes to the body's set of basic instructions, it raises many unique ethical concerns. The ethical questions surrounding gene therapy include:

- How can "good" and "bad" uses of gene therapy be distinguished?
- Who decides which traits are normal and which constitute a disability or disorder?
- Will the high costs of gene therapy make it available only to the wealthy?
- Could the widespread use of gene therapy make society less accepting of people who are different?
- Should people be allowed to use gene therapy to enhance basic human traits such as height, intelligence, or athletic ability?

Current gene therapy research has focused on treating individuals by targeting the therapy to body cells such as bone marrow or blood cells. This type of gene therapy cannot be passed on to a person's children. Gene therapy could be targeted to egg and sperm cells (germ cells), however, which would allow the inserted gene to be passed on to future generations. This approach is known as germline gene therapy.

The idea of germline gene therapy [7] is controversial. While it could spare future generations in a family from having a particular genetic disorder, it might affect the development of a fetus in unexpected ways or have long-term side effects that are not yet known. Because people who would be affected by germline gene therapy are not yet born, they can't choose whether to have the treatment. Because of these ethical concerns, the U.S. Government does not allow federal funds to be used for research on germline gene therapy in people.

Is gene therapy available to treat my disorder?

Gene therapy is currently available only in a research setting. The U.S. Food and Drug Administration (FDA) has not yet approved any gene therapy products for sale in the United States. Hundreds of research studies (clinical trials) are under way to test gene therapy as a treatment for genetic conditions, cancer, and HIV/AIDS. If you are interested in participating in a clinical trial, talk with your doctor or a genetics professional about how to participate.

You can also search for clinical trials online. ClinicalTrials.gov, a service of the National Institutes of Health, provides easy access to information on clinical trials. You can search for specific trials or browse by condition or trial sponsor. You may wish to refer to a list of gene therapy trials that are accepting (or will accept) participants.

This section [98], summarizes DNA-based therapeutics, RNA-based therapeutics, and gene transfer technologies for the diseases that are known to be genetic in origin. DNA-based therapeutics includes plasmids, oligonucleotides for antisense and antigene applications, DNA aptamers, and DNAzymes, while RNAbased therapeutics includes antisense RNA, ribozymes, RNA decoys, RNA aptamers, small interfering RNA, and microRNA. This section also includes the current status of gene therapy and recent developments in gene therapy research.

DNA-Based Therapeutics

Plasmids

Plasmids are high molecular weight, double-stranded DNA constructs containing transgenes, which encode specific proteins. On the molecular level, plasmid DNA molecules can be considered prodrugs that upon cellular internalization employ the DNA transcription and translation apparatus in the cell to biosynthesize the therapeutic entity, the protein [109,110]. The mechanism of action of plasmid DNA requires that the plasmid molecules gain access into the nucleus after entering the cytoplasm. Nuclear access or lack thereof eventually controls the efficiency of gene expression.

Oligonucleotides for Antisense and Antigene Applications

Oligonucleotides are short single-stranded segments of DNA that upon cellular internalization can selectively inhibit the expression of a single protein. For antisense applications, oligonucleotides interact and form a duplex with the mRNA or the pre-mRNA and inhibit its translation or processing, consequently inhibiting protein biosynthesis. For antigene applications, oligonucleotides must enter the cell nucleus, form a triplex with the double-stranded genomic DNA, and inhibit the translation as well as the transcription process of the protein. On the molecular level, numerous mechanisms have been proposed to explain the basis of oligonucleotide action [102,111].

Aptamers

DNA aptamers are double-stranded nucleic acid segments that can directly interact with proteins [112]. Aptamers interfere with the molecular functions of disease-implicated proteins or those that participate in the transcription or translation processes. Aptamers are preferred over antibodies in protein inhibition owing to their specificity, nonimmunogenicity, and stability of pharmaceutical

formulation [113]. DNA aptamers have demonstrated promise in intervention of pathogenic protein biosynthesis against HIV-1 integrase enzyme [114].

DNA zymes

DNA zymes are analogs of ribozymes with greater biological stability [115]. The RNA backbone chemistry is replaced by the DNA motifs that confer improved biological stability. DNA zyme directed against vascular endothelial growth factor receptor 2 was confirmed to be capable of tumor suppression by blocking angiogenesis upon intratumoral injections in mice [116].

RNA-Based Therapeutics

RNA Aptamers

RNA aptamers are single-stranded nucleic acid segments that can directly interact with proteins [112]. Aptamers recognize their targets on the basis of shape complementarity [117]. Moreover, their binding specificity and affinity for the target are extremely high and similar to monoclonal antibodies. RNA aptamers have demonstrated promise in intervention of pathogenic protein biosynthesis against HIV-1 transcriptase [118]. Moreover, RNA aptamers that specifically bind and inactivate vascular endothelial growth factor (VEGF) *in vitro* have been isolated.

RNA Decoys

The RNA decoys are designed to provide alternate, competing binding sites for proteins that act as translational activators or mRNA-stabilizing elements [119,120]. Decoys can prevent translation or induce instability and, ultimately, destruction of the mRNA. Over-expressed short RNA molecules corresponding to critical cis-acting regulatory elements can be used as decoys for trans-activating proteins, thus preventing binding of these Trans activators to their corresponding cis-acting elements in the viral genome [121,122]. RNA decoys have an advantage over other nucleic acid-based strategies, that is, the decoys are less likely to be affected by variability of the infectious agent because any mutation in the trans-activating protein affects not only binding to the decoys but also binding to the endogenous targets.

Antisense RNA

Antisense drugs are short stretches of deoxyribonucleotide analogs that bind to specific complementary areas of the mRNA by Watson–Crick base pairing to block gene expression in a sequence-specific fashion. The antisense drugs work at an early stage in the production of a disease-causing protein and theoretically can be applied to a number of diseases where the basic pathophysiology involves an overexpression of a given protein molecule [123,124]. Antisense ODNs can base pair with a gene's transcript and constitute a new technology for the control of gene expression in prokaryotes and eukaryotes, including mammalian cells [125].

A stoichiometric disadvantage of antisense RNA is the high expression required to successfully bind to all target RNA. A major advantage is the lack of immunogenicity of antisense constructs, such that the oligonucleotides and the cells producing them will not be destroyed by the host immune response. The antisense approach to modulating gene expression has been extensively reviewed [126].

Ribozymes

Antisense RNA alone is not potent enough to produce complete inhibition in vivo. An enzymatic moiety can be included with antisense oligonucleotide, which will cleave the target RNA once the RNA–RNA duplex has formed. These enzymatic RNA strands are called "Ribozymes." They are antisense RNA molecules that are capable of sequence-specific cleaving of RNA molecules [112]. They function by binding to the target moiety through antisense sequence-specific hybridization and inactivating it by cleaving the phosphodiester backbone at a specific site. Thus, they can selectively bind to target mRNAs and form a duplex having highly distorted confirmation that is easily hydrolyzed, and this hydrolysis of mRNA may be used for targeted suppression of specific gene [102].

Small Interfering RNAs

RNA interference (RNAi) is a posttranscriptional mechanism of gene silencing through chromatin remodeling, inhibition of protein translation, or direct mRNA degradation, which is ubiquitous in eukaryotic cells [127-129]. Small interfering RNAs (siRNAs) can be used for downregulation of disease-causing genes through RNA interference. Typically, these are short double-stranded RNA segments with 21–23 nucleotides and are complementary to the mRNA sequence of the protein whose transcription is to be blocked. On administration, siRNA molecules are incorporated into RNA-induced silencing complexes (RISC), which bind to the mRNA of interest and stimulate mRNA degradation mechanisms, such as nuclease activity, that lead to silencing of the particular gene [130-132].

Micro RNA

MicroRNAs (miRNAs) are a class of naturally occurring, small noncoding RNA molecules 21–25 nucleotides in length. These molecules are partially complementary to messenger RNA (mRNA) molecules, and their main function isdownregulation of gene expression via translational repression, mRNA cleavage, and deadenylation.

Gene Transfer Technologies

Gene transfer technologies can be classified into three general types: electrical techniques, mechanical transfection, and vector-assisted delivery systems.

Mechanical and Electrical Techniques

Strategies of introducing naked DNA into cells by mechanical and electrical techniques include microinjection, particle bombardment, the use of pressure, and electroporation. Microinjection is highly efficient since one cell at a time is targeted for DNA transfer, but it is time consuming. Ballistic transfer of gold microparticles may be performed using particle bombardment equipment such as the gene gun. Electroporation is achieved using high-voltage electrical current to facilitate DNA transfer that results in high cell mortality and is not suitable for clinical use [133-136].

Vector-Assisted Delivery Systems

Vector-assisted DNA/gene delivery systems can be classified into two types based on their origin: biological viral DNA delivery systems and chemical nonviral delivery systems.

Viral Delivery Systems

These viral DNA-delivery vectors include both RNA and DNA viruses. The viruses used as gene therapy vectors can be classified into four types: Retroviruses [137], Adenoviruses, Adeno-associated viruses [138], and Herpes simplex viruses. Gene expression using viral vectors has been achieved in tissues such as kidney [139], heart muscle [140], eye [138], and ovary [141]. Moreover, gene therapy using viral systems has made considerable progress for the treatment of a wide range of diseases, such as muscular dystrophy [140], AIDS [142], and cancer [143].

Nonviral Delivery Systems

Commonly used nonviral vectors for delivery of DNA-based therapeutics can be classified into three major types: Naked DNA delivery systems, polymeric delivery systems, and liposomal delivery systems [102, 144-146]. Recently, polyethylenimine (PEI), poly(lactic-co-glycolic acid) (PLGA), polypeptides, chitosan, cyclodextrin, dendrimers, and polymers containing different nanoparticles are used *in vitro* and *in vivo* with respect to their structure, physicochemical properties, and delivery efficiency as an siRNA delivery vehicle [147].

Recent Developments in Gene Therapy

Gene therapy is defined as the introduction of genetic material in a patient's cells with resulting therapeutic benefit. A promising new biomedical discipline could potentially lead to new treatments for hereditary diseases, cardiovascular and neurologic disorders, cancer, diabetes and even infectious diseases. The introduction of genetic material into somatic cells requires gene delivery vectors. Since viruses have developed, efficient means to introduce their own genetic material into cells they can be readily adapted as viral vectors for gene therapy.

Preclinical studies in animal models have shown that therapeutic effects can be achieved after gene therapy for genetic, acquired and complex disorders. Furthermore, therapeutic effects have been obtained in several phase I/II gene therapy clinical trials for hemophilia, severe combined immune deficiency (SCID) and cancer. Gene transfer technology has improved significantly over the past few years and has led to the development of vectors which have fewer side-effects without compromising their efficacy, at least partly due the development of cell-type specific targetable vectors. Nevertheless, the success of gene therapy is still very much depending upon the continuous development of improved vector technologies that would hopefully and ultimately cure diseases, which are refractory to current treatment paradigms. Status of gene therapy research and recent developments in gene therapy are given below: http://europepmc.org/search?query=AUTH:%22Vandendriessche+T%22&page=1&sortby=Date%2BDESC

1. Failed Downregulation of Circulating MicroRNA-155 in the Early Phase after ST Elevation Myocardial Infarction Is Associated with Adverse Left Ventricular Remodeling. (PMID:28618412) Latet SC, Van Herck PL, Claeys MJ, Van Craenenbroeck AH, Haine SE, Vandendriessche TR, Van Hoof VO, Fransen E, De Winter BY, Van Craenenbroeck EM, Heidbuchel H, Vrints CJ, Hoymans VY Cardiology [16 Jun 2017, 138(2):91-96]

2. AAV9 delivered bispecific nanobody attenuates amyloid burden in the gelsolin amyloidosis mouse model. (PMID:28605435) Verhelle A, Nair N, Everaert I, Van Overbeke W, Supply L, Zwaenepoel O, Peleman C, Van Dorpe J, Lahoutte T, Devoogdt N, Derave W, Chuah MK, VandenDriessche T, Gettemans J Hum Mol Genet [09 Jun 2017]

3. A Novel Platform for Immune Tolerance Induction in Hemophilia A Mice. (PMID:28552407) Merlin S, Cannizzo ES, Borroni E, Bruscaggin V, Schinco P, Tulalamba W, Chuah MK, Arruda VR, VandenDriessche T, Prat M, Valente G, Follenzi A Mol Ther [26 May 2017, 25(8):1815-1830]

4. AAV Capsid Engineering: Zooming in on the Target. (PMID:28498776) VandenDriessche T Hum Gene Ther [01 May 2017, 28(5):373-374]

5. Baboon envelope pseudotyped lentiviral vectors efficiently transduce human B cells and allow active factor IX B cell secretion in vivo in NOD/SCIDγc-/- mice. (PMID:27685947) Levy C, Fusil F, Amirache F, Costa C, Girard-Gagnepain A, Negre D, Bernadin O, Garaulet G, Rodriguez A, Nair N, Vandendriessche T, Chuah M, Cosset FL, Verhoeyen E J Thromb Haemost [29 Sep 2016, 14(12):2478-2492]

6. CRISPR/Cas9 Flexes Its Muscles: In Vivo Somatic Gene Editing for Muscular Dystrophy. (PMID:26952918 PMCID:PMC4786932) Free full text article VandenDriessche T , Chuah MK Mol Ther [01 Mar 2016, 24(3):414-416]

7. PiggyBac transposons expressing full-length human dystrophin enable genetic correction of dystrophic mesoangioblasts. (PMID:26682797 PMCID:PMC4737162) Free full text article Loperfido M, Jarmin S, Dastidar S, Di Matteo M, Perini I, Moore M, Nair N, Samara-Kuko E, Athanasopoulos T, Tedesco FS, Dickson G, Sampaolesi M, VandenDriessche T, Chuah MK Nucleic Acids Res [17 Dec 2015, 44(2):744-760]

8. Adeno-Associated Virus Type 2 and Hepatocellular Carcinoma? (PMID:26690810 PMCID:PMC4809064) Free full text article Berns KI, Byrne BJ, Flotte TR, Gao G, Hauswirth WW, Herzog RW, Muzyczka N, VandenDriessche T, Xiao X, Zolotukhin S, Srivastava A Hum Gene Ther [01 Dec 2015, 26(12):779-781]

9. Corrigendum: Efficient derivation and inducible differentiation of expandable skeletal myogenic cells from human ES and patient-specific iPS cells. (PMID:26313482) Maffioletti SM, Gerli MF, Ragazzi M, Dastidar S, Benedetti S, Loperfido M, VandenDriessche T, Chuah MK, Tedesco FS Nat Protoc [27 Aug 2015, 10(9):1457]

10. Gene therapy for cardiovascular disease: advances in vector development, targeting, and delivery for clinical translation. (PMID: 26239654 PMCID: PMC4571836) Free full text article Rincon MY, VandenDriessche T, Chuah MK Cardiovasc Res [03 Aug 2015, 108(1):4-20]

11. Pluripotent Stem Cells for Gene Therapy of Degenerative Muscle Diseases. (PMID:26122099) Loperfido M, Steele-Stallard HB, Tedesco FS, VandenDriessche T Curr Gene Ther [30 Jun 2015, 15(4):364-380]

12. Efficient derivation and inducible differentiation of expandable skeletal myogenic cells from human ES and patient-specific iPS cells. (PMID:26042384) Maffioletti SM, Gerli MF, Ragazzi M, Dastidar S, Benedetti S, Loperfido M, VandenDriessche T, Chuah MK, Tedesco FS Nat Protoc [04 Jun 2015, 10(7):941-958]

13. Moving forward toward a cure for hemophilia B. (PMID: 25943496 PMCID: PMC4427887) Free full text article VandenDriessche T, Chuah MK Mol Ther [01 May 2015, 23(5):809-811]

14. Usefulness of early rule-in and rule-out biomarker protocols to estimate ischemia-induced myocardial injury in early chest pain presenters. (PMID:25929579) Vorlat A, Van Hoof VO, Hammami R, van Kerckhoven S, Van der Heijden CM, Coenen D, Bosmans JM, Haine S, Vandendriessche TR, Vrints CJ, Claeys MJ Am J Cardiol [23 Mar 2015, 115(12):1667-1671]

15. Liver-directed lentiviral gene therapy in a dog model of hemophilia B.(PMID:25739762) Cantore A , Ranzani M , Bartholomae CC, Volpin M, Valle PD, Sanvito F, Sergi LS, Gallina P, Benedicenti F, Bellinger D, Raymer R , Merricks E, Bellintani F, Martin S, Doglioni C, D'Angelo A, VandenDriessche T, Chuah MK , Schmidt M, ..., Naldini L Sci Transl Med [01 Mar 2015, 7(277):277ra28]

16. Adiponectin and ischemia-reperfusion injury in ST segment elevation myocardial infarction. (PMID:25722457) De Roeck L, Vandamme S, Everaert BR, Hoymans V, Haine S, Vandendriessche T, Bosmans J , Ronsyn MW, Miljoen H, Van Berendoncks A, De Meyer G, Vrints C, Claeys MJ Eur Heart J Acute Cardiovasc Care [26 Feb 2015, 5(1):71-76]

17. Left ventricular remodeling following percutaneous edge-to-edge repair using Mitraclip. (PMCID:PMC4328177) Free full text article Paelinck BP, Vandendriessche T, De Bock D, Parizel PM, Rodrigus IE, Marc CJ J Cardiovasc Magn Reson [03 Feb 2015, 17(suppl 1):P343-P343]

18. Hitting the target without pulling the trigger. (PMID:25561034 PMCID:PMC4426814) Free full text article VandenDriessche T, Chuah MK Mol Ther [01 Jan 2015, 23(1):4-6]

19. Genome-wide computational analysis reveals cardiomyocyte-specific transcriptional Cis-regulatory motifs that enable efficient cardiac gene therapy. (PMID:25195597 PMCID:PMC4426801) Free full text article Rincon MY , Sarcar S , Danso-Abeam D , Keyaerts M , Matrai J , Samara-Kuko E , Acosta-Sanchez A , Athanasopoulos T , Dickson G , Lahoutte T , De Bleser P , VandenDriessche T , Chuah MK Mol Ther [08 Sep 2014, 23(1):43-52]

20. Hyperactive PiggyBac Transposons for Sustained and Robust Liver-targeted Gene Therapy. (PMID:28153205) Di Matteo M , Samara-Kuko E , Ward NJ , Waddingon SN , McVey JH , Chuah MK , VandenDriessche T Mol Ther [01 Sep 2014, 22(9):1614-1624]

21. Treatment of phenylketonuria using minicircle-based naked-DNA gene transfer to murine liver. (PMID:24585515 PMCID:PMC4449723) Free full text article Viecelli HM , Harbottle RP , Wong SP , Schlegel A , Chuah MK , VandenDriessche T , Harding CO , Thöny B Hepatology [29 Jul 2014, 60(3):1035-1043]

22. Hyperactive piggyBac transposons for sustained and robust liver-targeted gene therapy. (PMID: 25034357 PMCID: PMC4435487) Free full text article Di Matteo M , Samara-Kuko E , Ward NJ , Waddington SN , McVey JH , Chuah MK , VandenDriessche T Mol Ther [18 Jul 2014, 22(9):1614-1624]

23. Liver-specific transcriptional modules identified by genome-wide in silico analysis enable efficient gene therapy in mice and non-human primates. (PMID: 24954473 PMCID: PMC4435486) Free full text article Chuah MK, Petrus I, De Bleser P, Le Guiner C, Gernoux G, Adjali O, Nair N, Willems J, Evens H, Rincon MY, Matrai J, Di Matteo M, Samara-Kuko E, Yan B , Acosta-Sanchez A, Meliani A, Cherel G, Blouin V, Christophe O, ..., VandenDriessche T Mol Ther [23 Jun 2014, 22(9):1605-1613]

24. Percutaneous mitral valve repair in high-risk patients: initial experience with the Mitraclip system in Belgium. (PMID: 25029871) Vandendriessche T, Kotrc M, Tijskens M , Bartunek J, Delesie M, Paelinck BP, De Bock D, Penicka M, Stockman B, De Maeyer C, Vrints C, Vanderheyden M, Claeys MJ Acta Cardiol [01 Jun 2014, 69(3):265-270]

25. Cost-effectiveness of contemporary vascular closure devices for the prevention of vascular complications after percutaneous coronary interventions in an all-comers PCI population.(PMID:24952056) Kerré S , Kustermans L, Vandendriessche T, Bosmans J, Haine SE, Miljoen H, Vrints CJ, Beutels P, Wouters K, Claeys MJ EuroIntervention [01 Jun 2014, 10(2):191-197]

Conclusions

- Nearly every cell in a person's body has the same DNA located in the cell nucleus.
- Most human diseases are complex multi-factorial diseases resulting from the combination of various genetic and environmental factors.
- A number of human diseases are known to be genetic in origin.
- The ability of DNA bases to form wobble base pairs as well as Watson-Crick base pairs can result in base-pair mismatches occurring during DNA replication. If not repaired by DNA repair enzymes, these mismatches can lead to genetic diseases and cancer.

- At present, there is no drug really able to protect DNA; the only mean available is the decrease of the exposure to the risk factors such as sun, tobacco, irradiations, and toxic compounds.
- Modern drug research aims to discover biologically active molecule(s) that are absolutely specific to the molecular targets responsible for the disease progression.
- "Nucleic acid medicine" is a next generation drug discovery technology with a completely different mechanism of action than traditional pharmaceutical products.
- In recent years nucleic acid-based drugs have emerged extremely promising candidates for drug therapy to a wide range of diseases, including cancer, infectious diseases, diabetes, cardiovascular, inflammatory, and neurodegenerative diseases, cystic fibrosis, hemophilia, and other genetic disorders.
- The number of atomic-level quadruplex structures from NMR and crystallography is still tantalizingly small, and provides insight into the finely balanced equilibria between different topologies in vitro.
- Gene therapy is a technique for correcting defective genes responsible for disease development, which may be classified into two types: somatic and germ line gene therapy.
- There are many ethical, social, and commercial issues raised by the prospects of treating patients using gene therapy.

References

1. Drugs and nucleic acids, Pharmacology, DNA and RNA biosynthesis, inhibitors of nucleotides, Oligonucleotides as drugs.
2. Minoru Kanehisa., *et al*. "KEGG for representation and analysis of molecular networks involving diseases and drugs". *Nucleic Acids Research* 38 (2010): 355-360.
3. International HapMap Consortium. "The International HapMap Project". *Nature* 426 (2003): 789-796.
4. Hirschhorn JN and Daly MJ. "Genome-wide association studies for common diseases and complex traits". *Nature Reviews Genetics* 6.2 (2005): 95-108.
5. Stratton MR., *et al.* "The cancer genome". *Nature* 458 (2009): 719-724.
6. Helen MBerman and Peter R Young. "Annual Reviews Biophs.Bioeng". 10 (1981): 87-114.
7. Lister Hill National Center for Biomedical Communications U.S. National Library of Medicine
8. National Institutes of Health Department of Health & Human Services, Published August 2, (2017).
9. Patrick and Corey. "Topic 5 Nucleic Acids as Drug Targets". *Nucleic Acids-Chapter* 7 187-188 193-194 198-199.
10. David S Wishart., *et al.* "DrugBank: a comprehensive resource for in silico drug discovery and exploration". *Nucleic Acids Research* 34 (2006): 668-672.
11. Chen X., *et al.* "TTD: therapeutic target database". *Nucleic Acids Research* 30.1 (2002): 412-415.
12. Kanehisa M., *et al*. "The KEGG resource for deciphering the genome". Nucleic Acids Research 32 (2004): 277-280.
13. Brooksbank C., *et al*. "The European Bioinformatics Institute's data resources: towards systems biology". *Nucleic Acids Research* 33 (2005): 46-53.
14. Nucleic Acids Book (www.atdbio.com/nucleic-acids-book).
15. Oganesian L and Bryan TM. "Physiological relevance of telomeric Gquadruplex formation: a potential drug target". *Bioessays* 29.2 (2007): 155-165.
16. De Cian A., *et al.* "Targeting telomeres and telomerase". *Biochimie* 90.1 (2008): 131-155.
17. Ou T., *et al*. "G-Quadruplexes: targets in anticancer drug design". *ChemMedChem* 3.5 (2008): 690-713.
18. See www.cylenepharma.com for a more detailed description.
19. Burge SE., *et al*. "Quadruplex DNA: sequence, topology and structure". *Nucleic Acids Research* 34.19 (2006): 5402-5415.
20. Patel DJ., *et al.* "Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics". *Nucleic Acids Research* 35.22 (2007): 7429-7455.
21. Engelhart AE., *et al.* "Metal ion interactions with G-quadruplex structures". *In Nucleic Acid–Metal Ion Interactions* 2008: 118-153.

22. Šˇponer J and Šˇpacˇkova´ N. "Molecular dynamics simulations and their application to four-stranded DNA". *Methods* 43.4 (2007): 278-290.

23. Pe´rez A., *et al*. "Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers". *Biophysical Journal* 92.11 (2007): 3817-3829.

24. Li H., *et al.* "Force-induced unfolding of human telomeric G-quadruplex: a steered molecular dynamics simulation study". *Biochemical and Biophysical Research Communications* 79.1 (2008): 70-75.

25. Zaug AJ., *et al*. "Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension *in vitro*". *Proceedings of the National Academy of Sciences* 102.31 (2005): 10864-10869.

26. Possemato R., *et al*. "Suppression of hPOT1 in diploid human cells results in an hTERT-dependent alteration of telomere length dynamics". *Molecular Cancer Research* 6.10 (2008): 1582-1593.

27. Sun D., *et al.* "Inhibition of human telomerase by a G-quadruplex-interactive compound". *Journal of Medicinal Chemistry* 40.14 (1997): 2113-2116.

28. Zahler AM., *et al*. "Inhibition of telomerase by G-quartet DNA structures". *Nature* 350.6320 (1991): 718-720.

29. Gomez D., *et al*. "The G-quadruplex ligand telomestatin inhibits pot1 binding to telomeric sequences in vitro and induces GFP-POT1 dissociation from telomeres in human cells". *Cancer Research* 66.14 (2006): 6908-6912.

30. Phatak P., *et al.* "Telomere uncapping by the Gquadruplex ligand RHPS4 inhibits clonogenic tumour cell growth *in vitro* and *in vivo* consistent with a cancer stem cell targeting mechanism". *British Journal of Cancer* 96.8 (2007): 1223-1233.

31. Gunaratnam M., *et al*. "Mechanism of acridine-based telomerase inhibition and telomere shortening". *Biochemical Pharmacology* 74.1 (2007): 679-689.

32. Salvati E., *et al*. "Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect". *Journal of Clinical Investigation* 117.11 (2007): 3236-3247.

33. Rodriquez R., *et al*. "A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres". *Journal of the American Chemical Society* 130.47 (2008): 15758-15759.

34. Ying L., *et al.* "Studies on the structure and dynamics of the human telomeric G quadruplex by single-molecule fluorescence resonance energy transfer". *Proceedings of the National Academy of Sciences of the United States of America* 100.25 (2003): 14629-14634.

35. Rujan IN., *et al*. "Vertebrate telomere repeat DNAs favor external loop propeller quadruplex structures in the presence of high concentrations of potassium". *Nucleic Acids Research* 33.6 (2005): 2022-2031.

36. Wang Y and Patel DJ. "Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex". *Structure* 1.4 (1993): 263-282.

37. Parkinson GN., *et al.* "Crystal structure of parallel quadruplexes from human telomeric DNA". *Nature* 417 (2002): 876-880.

38. Luu KN., *et al.* "Structure of the human telomere in K+ solution: an intramolecular (3 + 1) Gquadruplex scaffold". *Journal of the American Chemical Society* 128.30 (2006): 9963-9970.

39. Ambrus A., *et al.* "Human telomeric sequence forms a hybrid-type intramolecular Gquadruplex structure with mixed parallel/antiparallel strands in potassium solution". *Nucleic Acids Research* 34.9 (2006): 2723-2735.

40. Phan AT., *et al.* "Different loop arrangements of intramolecular human telomeric (3 + 1) G-quadruplexes in K+ solution". *Nucleic Acids Research* 34.19 (2006): 5715-5719.

41. Dai J., *et al*. "Structure of the hybrid-2 type intramolecular human telomeric G-quadruplex in K+ solution: insights into structure polymorphism of the human telomeric sequence". *Nucleic Acids Research* 35.15 (2007): 4927-4940.

42. Phan AT., *et al*. "Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter". *Journal of the American Chemical Society* 129.14 (2007): 386-4392.

43. Schoeftner S and Blasco MA. "Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II". *Nature Cell Biology* 10 (2008): 228-236.

44. Azzalin CM., *et al*. "Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends". *Science* 318.5851 (2007): 798-801.

45. Xu Y., *et al*. "Human telomeric RNA in G-quadruplex structure". *Nucleic Acids Sympos Ser* 52 (2008): 175-176.

46. Xu Y., *et al*. "G-quadruplex formation by human telomeric repeats-containing RNA in Na+ solution". *Journal of the American Chemical Society* 130.33 (2008): 1179-11184.

47. Monchaud D and Teulade-Fichou MP. "A hitchhiker's guide to G-quadruplex ligands". *Organic and Biomolecular Chemistry* 6.4 (2008): 627-636.

48. Clark GR., *et al*. "Structure of the first parallel DNA quadruplex-drug complex". *Journal of the American Chemical Society* 125.14 (2003): 4066-4067.

49. Campbell NH., *et al*. "Structural basis of DNA quadruplex recognition by an acridine drug". *Journal of the American Chemical Society* 130.21 (2008): 6722-6724.

50. Haider SM., *et al*. "Structure of a G-quadruplex–ligand complex". *Journal of Molecular Biology* 326.1 (2003): 117-125.

51. Gavathiotis E., *et al.* "Drug recognition and stabilisation of the parallel-stranded DNA quadruplex d (TTAGGGT) 4 containing the human telomeric repeat". *Journal of Molecular Biology* 334.1 (2003): 25-36.

52. Phan AT., *et al*. "Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter". *Nature Chemical Biology* 1.3 (2005): 167-173.

53. Parkinson GN., *et al.* "Structural basis for binding of porphyrin to human telomeres". *Biochemistry* 46.9 (2007): 2390-2397.

54. Parkinson GN., *et al.* "Topology conservation and loop flexibility in quadruplex-drug recognition: crystal structures of inter- and intramolecular telomeric DNA quadruplex-drug complexes". *Journal of Molecular Biology* 381.5 (2008): 1145-1156.

55. Huppert JL and Balasubramanian S. "Prevalence of quadruplexes in the human genome". *Nucleic Acids Research* 33.9 (2005): 2908-2916.

56. Todd AK., *et al*. "Highly prevalent putative quadruplex sequence motifs in human DNA". *Nucleic Acids Research* 33.9 (2005): 2901-2907.

57. The very useful web site www.quadruplex.org has links to various quadruplex databases and search engines, as well as much other useful information on quadruplexes and their literature.

58. Hershman SG., *et al*. "Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in Saccharomyces cerevisiae". *Nucleic Acids Research* 36.1 (2008): 144-156.

59. Verma A., *et al*. "Genome-wide computational and expression analyses reveal G-quadruplex DNA motifs as conserved cis-regulatory elements in human and related species". *Journal of Medicinal Chemistry* 51.18 (2008): 5641-5649.

60. Rawal P., et al. "Genome-wide prediction of G4 DNA as regulatory motifs: role in Escherichia coli global regulation". *Genome Research* 16.5 (2006): 644-655.

61. Huppert JL and Balasubramanian S. "G-quadruplexes in promoters throughout the human genome". Nucleic Acids Research 35.2 (2007): 406-413.

62. Du Z., *et al*. "Genome-wide analysis reveals regulatory role of G4 DNA in gene transcription". Genome Research 18.2 (2008): 233-241.

63. Huppert JL., *et al*. "Gquadruplexes: the beginning and end of UTRs". *Nucleic Acids Research* 36.19 (2008): 6260-6268.

64. Hazel P., *et al*. "Loop-lengthdependent folding of G-quadruplexes". *Journal of the American Chemical Society* 126.50 (2004): 16405-16415.

65. Risitano A and Fox KR: "Influence of loop size on the stability of intramolecular DNA quadruplexes". *Nucleic Acids Research* 32.8 (2004): 2598-2606.

66. Bugaut A and Balasubramanian S. "A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadruplexes". *Biochemistry* 47.2 (2008): 689-697.

67. Siddiqui Jain A., *et al*. "Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription". *Proceedings of the National Academy of Sciences* 99.18 (2002): 11593-11598.

68. Stephen Neidle. "The structures of quadruplex nucleic acids and their drug complexes". *Current Opinion in Structural Biology* 19.3 (2009): 239–250.

69. R. Diamond. In Computational Crystallography, D. Sayre, Ed. (Oxford Univ. Press, Oxford) T. A. Jones, in ibid 1982: 318-303.

70. FP Brooks Jr. in Proceedings of the 1977 International Federation of Information Proc-essing, B. Gilchrist, Ed. (North-Holland, Am-sterdam), 1977: 625.

71. CD Barry, CE Molnar, FU Rosenberger, Technical Memo 229 (Computer Systems Labo-ratory, Washington University, St. Louis, Mo., January 1976); C. D. Barry, H. E. Bosshard, R. A. Ellis, G. R. Marshall, in Computers in Life Science Research, W. Siler and D. A. B. Lind-berg, Eds. (Plenum, New York, 1975), p. 137.

72. TP orter, Computers & Graphics 12, 282 (1978); ibid. 13, 234 (1979).

73. RJ Feldmann., *et al*. Proceedings of the National Academy of Sciences of the United States America 75.5409 (1978).

74. K Knowlton and L. Cherry, Comput. Chem. 1 161 (1977).

75. NL Max, Computers & Graphics 13 165 (1979).

76. CK Johnson, Oak Ridge Natl. Lab. Tech. Rep. 5138 (1976).

77. FM Richards, Annual Review of Biophysics and Bioengineering 6 151 (1977).

78. B Lee and FM Richards, Journal of Molecular Biology 55 379 (1971).

79. A Shrake and JA Rupley, Journal of Molecular Biology 79 351 (1973).

80. ML Connolly, thesis, University of California, Berkeley (1981).

81. QCPE Bull. The dot molecular surface program (MS) is written in FORTRAN and may be obtained by writing to Quantum Chemis-try Program Exchange, De-partment of Chemistry, Indiana University, Bloomington (1981): 47405-47475.

82. R Huber, D Kukla, W Bode, P Schwager, K Bartels, J Deisenhofer, W Steigemann. Journal of Molecular Biology (1974): 73-89.

83. RN Smith, C Hansch, FH Kim, B Omiya, G Fukumura, CD Selassie, PYC Jow, JM Blaney, R Langridge, Arch. Biochemical and Biophysi-cal Research Communications (1982): 215-319.

84. C Hansch, R Li, JM Blaney, R Langridge, Journal of Medicinal Chemistry (1982): 25-777.

85. S Sprang, R Fletterick, M Stern, D Yang, N Madsen, J. Sturtevant, Biochemistry 21 2036 (1982).

86. E Goldsmith, S Sprang, R Fletterick, Journal of Molecular Biology 156 411 (1982).

87. SR Sprang, EJ Goldsmith, RJ Fletterick, SG Withers, NB Madsen, Biochemistry 21 5364 (1982).

88. RA Lerner, Nature (London) 299 592 (1982).

89. AJ Olson, G Cohen, D Davies, Antibody Structure, film available from Byron Motion Pictures, 65 K Street, NE, Washington, D.C. 20002.

90. J. Olson, Virus Wars, computer-generated film presented at the International School of Crystallography, Conference on Crystal-logra-phy in Molecular Biology, Italy, June 1982.

91. RE Dickerson and I Geis. "Hemoglobin: Structure, Function, Evolution, and Pathology". (Benjamin/Cummings, Menlo Park, Calif) (1983): 136.

92. IDKuntz, J M Blaney, S J Oatley, R Langridge, TE Ferrin. Journal of Molecular Biology 161 269 (1982).

93. R Langridge, TE Ferrin, ID Kuntz, ML Connolly. Science 211 661 (1981).

94. RO Fox Jr and FM Richards. Nature (London) 300 325 (1982).

95. Michael L Connolly. "Solvent-Accessible Surfaces of Proteins and Nucleic Acids". *Science* 221.4612 (1983): 709-713.

96. Solid-phase oligonucleotide synthesis

97. www.atdbio.com/content/17/Solid-phase-oligonucleotide-synthesis

98. (www.atdbio.com/nucleic-acids-book)

99. http://www.bonac.com/global/en/nucleic/

100. Nucleic acid medicine: Next generation drugs using DNA and RNA; with research conducted in countries around the world.

101. Current Medicinal Chemistry 20.28 (2013): 3515-38.

102. Saraswat Pushpendra., *et al.* "From Nucleic Acids Sequences to Molecular Medicine". RNA Technologies (2012).

103. Mizutani TN., et al. "Inhibition of hepatitis C virus replication by antisense oligonucleotids in culture cells". *Biochemical and Bio-physical Research Communications* 212.3 (1995): 906–911.

104. Ehsan A., *et al.* "Long-term stabilization of vein graft wall architecture and prolonged resistance to experimental atherosclerosis after E2F decoy oligonucleotide gene therapy". *Journal of Thoracic and Cardiovascular Surgery* 121.4 (2001): 714-722.

105. Patil SD and Burgess DJ. "DNA-based pharmaceuticals: therapeutics for the 21st century". *American Association of Pharmaceutical Scientists* 6.27 (2003).

106. Mardan T., *et al.* "Prospects for cationic polymers in gene and oligonucleotide therapy against cancer". *Advanced Drug Delivery Reviews* 54.5 (2002): 715-758.

107. Shi N., *et al.* "Brain-specific expression of an exogenous gene after IV administration". *Proceedings of the National Academy of Sciences of the United States of America* 98.22 (2001): 12754-12759.

108. SoRelle R. "Who owns your DNA? Who will own it"? *Circulation* 101 (2000):67-68.

109. "Guidance on making proposals to conduct gene therapy research on human subjects. Report of the United Kingdom Health Ministers' Gene Therapy Advisory Committee". *Human Gene Therapy* 6.3 (1995): 335-346.

110. Baker BF. "The role of antisense oligonucleotides in the wave of genomic information". *Nucleosides Nucleotides Nucleic Acids* 20.4.7 (2001): 397-399.

111. Van Ommen GJB., *et al.* "the human genome project and the future of diagnostics, treatment, and prevention". *Lancet* 354 (1999): 5-10.

112. Rosenberg SA., *et al.* "Human gene marker/therapy protocols". *Human Gene Therapy* 10.18 (1999): 3067-3123.

113. Uherek C and Wels W. "DNA-carrier proteins for targeted gene delivery". *Advanced Drug Delivery Reviews* 44.2.3 (2000):153-166.

114. Zhaohui P. "China OKs gene therapy drug". *Genetic Engineering News* 6 (2003): 53.

115. Speedie M. Antisense therapeutic agents. In: Williams DA, Lemke TL (eds) Foye's principles of medicinal chemistry, 5th edn. Lippincott, Philadelphia, PA. (2005).

116. Stull RA and Szoka FC Jr. "Antigene, ribozyme and aptamer nucleic acid drugs: progress and Prospects". *Pharmaceutical Research* 12.4 (1995): 463-465.

117. Jayasena SD. "Aptamers: an emerging class of molecules that rival antibodies in diagnostics". *Clinical Chemistry* 45.9 (1999): 1628-1650.

118. De Saultrait VR., et al. "DNA aptamers derived from HIV-1 RNaseH inhibitors are strong anti-integrase agents". *Journal of Molecular Biology* 324.2 (2002): 195-203.

119. Akhtar S., *et al.* "The delivery of antisense therapeutics". *Advanced Drug Delivery Reviews* 44.1 (2000): 3-21.

120. Zhang L., *et al.* "Angiogenic inhibition mediated by a DNAzyme that target vascular endothelial growth factor receptor 2". *Cancer Research* 62.19 (2002): 5463-5469.

121. Kaur G and Roy I. "Therapeutic applications of aptamers". *Expert Opinion on Investigational Drugs* 17.1 (2008): 43-60.

122. Chaloin L., *et al.* "Endogenous expression of a high-affinity pseudoknot RNA aptamer suppresses replication of HIV-1". *Nucleic Acids Research* 30.18 (2002): 4001-4008.

123. Beelman CA and Parker R. "Degradation of mRNA in eukaryotes". *Cell* 81.2 (1995): 179-183.

124. Liebhaber SA. "mRNA stability and the control of gene expression". *Nucleic Acids Symposium Series* 36 (1997): 29-32.

125. Sullenger BA., *et al.* "Overexpression of TAR sequences renders cells resistant to human immunodeficiency virus replication". *Cell* 63.3 (1990): 601-608.

126. Sullenger BA*., et al.* "Analysis of trans-acting response decoy RNAmediated inhibition of human immunodeficiency virus type 1 transactivation". *Journal of Virology* 65.12 (1991): 6811-6816.

127. Crooke ST. "an overview of progress in antisense therapeutics". *Antisense & Nucleic Acid Drug Development* 8.2 (1998): 115-122.

128. Speedie M. "Antisense therapeutic agents". In: Williams DA, Lemke TL (eds) Foye's principles of medicinal chemistry, 5th edn. Lippincott, Philadelphia, PA. (2005).

129. Zamecnik PC and Stephenson ML. "Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotides". *Proceedings of the National Academy of Sciences of the United States of America* 75.1 (1978): 280-284.

130. Gewirtz AM., *et al.* "Nucleic acid therapeutics: state of the art and future prospects". *Blood* 92.3 (1998): 712-736.

131. Caplen NJ. "Gene therapy progress and prospects. Downregulating gene expression: the impact of RNA interference". *Gene Therapy* 11.16 (2004): 1241-1248.

132. Dorsett Y and Tuschl T. "siRNAs: applications in functional genomics and potential as therapeutics". *Nature Reviews Drug Discovery* 3.4 (2004): 318-329.

133. Shankar P., *et al*. "The prospect of silencing disease using RNA interference". *Journal of the American Medical Association* 293.11 (2005): 1367-1373.

134. Bertrand JR., et al. "Comparison of antisense oligonucleotides and siRNAs in cell culture and *in vivo*". *Biochemical and Biophysical Research Communications* 296.4 (2002): 1000-1004.

135. McMinus MT and Sharp PA. "Gene silencing in mammals by small interfering RNAs". *Nature Reviews Genetics* 3.10 (2002): 737-747.

136. Scherr M., *et al.* "Gene silencing mediated by small interfering RNAs in mammalian cells". *Current Medicinal Chemistry* 10.3 (2003): 245-256.

137. McAlister DV., *et al.* "Microfabricated microneedles for gene and drug delivery". *Annual Review of Biomedical Engineering* 2 (2000): 289-313.

138. Luo D and Saltzman WM. "Synthetic DNA delivery systems". Nature Biotechnology 18.1 (2000): 33-37.

139. Regnier V., et al. "Electroporation-mediated delivery of 30-protected phosphodiester oligodeoxynucleotides to the skin". *Journal of Controlled Releas* 67 (2000): 337-346.

140. Huang L, Viroonchatapan E. Introduction. In: Huang L, Huang M-C, Wagner E (eds) Nonviral vectors for gene therapy. Academic Press, San Diego, CA, (1999): 3–22.

141. McTaggart S and Al-Rubeai M. "Retroviral vectors for human gene delivery". *Biotechnology Advances* 20.1 (2002): 1-31.

142. Martin KR., *et al.* "Gene delivery to the eye using adeno-associated viral vectors". *Methods* 28.2 (2002): 267-275.

143. Lien Y-HH and Lie L-W. "Gene therapy for renal disorders: what are the benefits for the elderly?" *Drugs & Aging* 19.8 (2002): 553-560.

144. Chamberlain JS. "Gene therapy of muscular dystrophy". *Human Molecular Genetics* 11.20 (2002): 2355–2362.

145. Wolf JK and Jenkins AD. "Gene therapy for ovarian cancer". *International Journal of Oncology* 21.3 (2002): 461-468.

146. Lever AML. "Gene therapy in the fight against AIDS". *Expert Opinion on Therapeutic Patents* 6 (1996): 161-167.

147. Zhao W., *et al*. "Adenoviral vectors for cancer gene therapy". *Current Genomics* 3.3 (2002): 163-180.

148. Fattal E., *et al*. "Liposomes for the delivery of nucleotides and oligonucleotides". *STP pharma sciences* 9.5 (1999): 383-390.

149. Fattal E., *et al.* "Liposome-based formulations for the delivery of oligonucleotids". *STP pharma sciences* 11.1 (2001): 31-44.

150. Pedroso de Lima MC., *et al.* "Cationic lipid–DNA complexes in gene delivery: from biophysics to biological applications". *Advanced Drug Delivery Reviews* 47.2.3 (2001): 277-294.

151. Kaushik S., *et al*. "Polymers in small-interfering RNA delivery". *Nucleic Acid Ther* 21.3 (2011): 133-147.