DNA TRANSCRIPTION 2024

Tuan Tran, M.D., Ph.D.

1. DNA transcription:

Transcription of DNA is the process of rewriting a segment of DNA into RNA.
Only **one** of the two DNA strands, called the template strand, is replicated during transcription. The antisense strand of DNA is read by RNA polymerase from the 3' end to the 5' end during transcription (3' → 5'). The complementary RNA is created in the opposite direction, in the 5' → 3' direction, matching the sequence of the sense strand except switching uracil for thymine. This directionality is because RNA polymerase can only add nucleotides to the 3' end of the growing mRNA chain. This use of only the 3' → 5' DNA strand eliminates the need for Okazaki fragments (seen in DNA replication) as well as that for an RNA primer to initiate RNA synthesis, as is the case in DNA replication. The resulting RNA strands are **single-stranded messenger RNA** (**mRNA**).

The overall process of transcription is a highly controlled process catalyzed and regulated by the enzyme DNA-dependent RNA polymerase.

- The first step is the recognition of specific DNA sequences called *promoter* sequences that signify the beginning of the gene.
- It is then followed by the separation of two strands of DNA and replication of **one** of the strands by the RNA polymerase.
- The sequence formed after replication is complementary to the template sequence as the process follows the complementary base pairing rules of DNA, except the thymine is replaced by uracil.

In eukaryotes, various proteins, such as the transcription factors (TFs), are involved in the regulation of transcription.

- Besides, the post-transcriptional modification also takes place in eukaryotes where the pre-mRNA (the result of transcription) is edited by the process of **splicing** before the mature mRNA reaches ribosomes for translation.
- The mRNA thus produced acts as a blueprint for protein synthesis during the process of translation.
- Depending on the sequence of DNA chosen for transcription, rRNA and tRNA synthesis might also happen.
- Transcription occurs in the nucleus of eukaryotes and cytoplasm of prokaryotes where the enzymes and the TFs are available.
- It may be inhibited by some antibiotics like rifampicin and 8-Hydroxyquinoline.
- The process can be detected by methods like Reverse Transcriptase PCR (RT-PCR), DNA microarray, in-situ hybridization, and Northern blotting.

2. Transcription Factors (TFs)

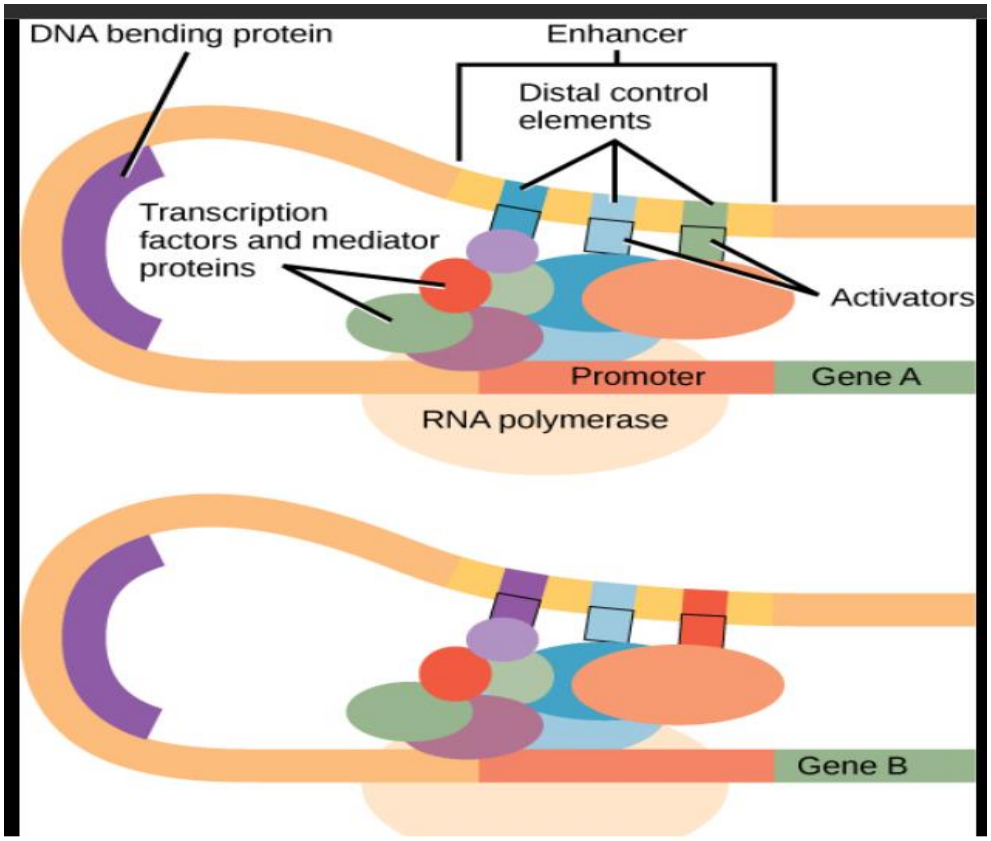
Transcription factors are proteins that play a crucial role in regulating gene expression. They bind to specific DNA sequences, either alone or in combination with other proteins, to control the rate of transcription of DNA to messenger RNA (mRNA). Here are some key points:

- **Function**: Transcription factors can act as activators or repressors. Activators increase the rate of transcription, while repressors decrease it.
- **Binding Sites**: They bind to specific DNA sequences known as enhancers or promoters.
- **Types**: There are ***general*** *transcription factors* (***G****TFs*), which are necessary for the transcription of all genes, and ***specific*** *transcription factors*, which regulate the transcription of particular genes.

Promoters

Promoters are specific DNA sequences located near the start of a gene. They serve as binding sites for transcription factors and RNA polymerase, the enzyme responsible for synthesizing RNA from the DNA template. Here are some key points:

- **Location**: Promoters are typically found upstream (5' end) of the gene they regulate.
- **Components**: A promoter region often contains several key elements, including the TATA box, which is a sequence of thymine (T) and adenine (A) nucleotides.
- **Function**: The primary role of promoters is to initiate transcription by providing a binding site for RNA polymerase and transcription factors.



LEGEND:

Transcription preinitiation complex (TPIC), represented by the central cluster of proteins, causes RNA polymerase to bind to target DNA site. The TPIC is able to bind both the promoter sequence near the gene to be transcribed and an enhancer sequence in a different part of the genome, allowing enhancer sequences to regulate a gene distant from it.

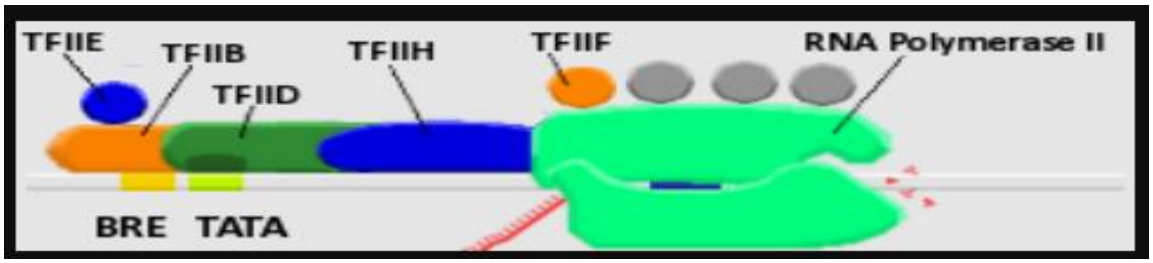
How they work together:

i. **Initiation**: Transcription factors first bind to the promoter region of a gene.

ii. **Recruitment**: This binding recruits RNA polymerase to the promoter.

iii. **Transcription Pre-Initiation Complex**: Transcription initiation requires an RNA polymerase II and a set of *multiple* General Transcription Factors (GTFs) to form a transcription preinitiation complex, which unwinds the DNA and begins synthesizing mRNA from the DNA template. The multiple GTFs are:

- TFIIA – stabilizes the interaction between the TATA box and TFIID/TATA binding protein (TBP)
- TFIIB – recognizes the *B recognition element* (BRE) in promoters
- TFIID – binds to TBP and recognizes TBP associated factors (TAFs), also adds promoter selectivity
- TFIIE – attracts and regulates TFIIH
- TFIIF – stabilizes RNA polymerase interaction with TBP and TFIIB; helps attract TFIIE and TFIIH
- TFIIH – unwinds DNA at the transcription start point, phosphorylates Ser5 of the RNA polymerase CCTD, releases RNA polymerase from the promoter



Understanding the roles of transcription factors and promoters is essential for grasping how genes are regulated and expressed in cells.

3. The Transcription Start Site. Its role in Transcription & Splicing:

The **Transcription Start Site (TSS)** is a fundamental genomic landmark that marks the point where RNA polymerase binds to the DNA and begins transcribing RNA. Located on the DNA template strand, the TSS marks the initiation point of transcription (where RNA polymerase binds to the DNA and begins transcribing RNA).

Role in Transcription

i. **Initiation of Transcription**: The TSS is the nucleotide position where transcription starts. RNA polymerase binds to the promoter region near the TSS and begins synthesizing RNA from the DNA template.

ii. **Promoter Region**: The promoter region, located upstream of the TSS, contains specific sequences that are recognized by transcription factors and RNA polymerase. These sequences help recruit the transcription machinery to the correct location.

iii. **Regulation of Gene Expression**: The precise location of the TSS is critical for the regulation of gene expression. Different TSSs can be used under different conditions, leading to the production of distinct mRNA isoforms with varying regulatory elements.

Role in Splicing

i. **Alternative Transcription Start Sites (ATSS)**: The use of alternative TSSs can influence splicing by generating different 5' untranslated regions (5' UTRs). These variations can affect the inclusion or exclusion of exons during splicing, thereby producing different mRNA isoforms.

ii. **Coordination with Splicing**: The choice of TSS can impact the splicing process by altering the context in which splicing occurs. For example, the presence of specific regulatory elements in the 5' UTR can influence the binding of splicing factors and the selection of splice sites.

iii. **Impact on mRNA Stability and Translation**: The different mRNA isoforms produced by alternative TSSs can have varying stabilities and translation efficiencies. This can further influence the overall gene expression and protein production.

Summary

The TSS is essential for the initiation of transcription and plays a significant role in regulating gene expression. Its position and usage can influence splicing patterns, leading to the production of diverse mRNA isoforms that contribute to cellular function and adaptation.

4. Promoters & TATA box:

4.1. Promoter of a Gene

-A promoter is a sequence of DNA located upstream (towards the 5' region of the sense strand) of the transcription start site (TSS) of a gene. It serves as the binding site for RNA polymerase and various transcription factors, initiating the process of transcription.

-Promoters can be about 100–1000 base pairs long, the sequence of which is highly dependent on the gene and product of transcription, type or class of RNA polymerase recruited to the site, and species of organism.

-Promoters control gene expression in bacteria and eukaryotes. RNA polymerase must attach to DNA near a gene for transcription to occur.

-A promoter typically includes several key elements:

i. **Core Promoter**: This includes the TSS and surrounding sequences, essential for the basic transcription machinery to bind and start transcription.

ii. **Proximal Promoter Elements**: These are located close to the TSS and include specific sequences like the TATA box, CAAT box, and GC-rich regions. They help recruit transcription factors and RNA polymerase.

iii. **Distal Promoter Elements**: These can be located further upstream and include enhancers and silencers that modulate the efficiency and rate of transcription[1].

4.2. TATA Box

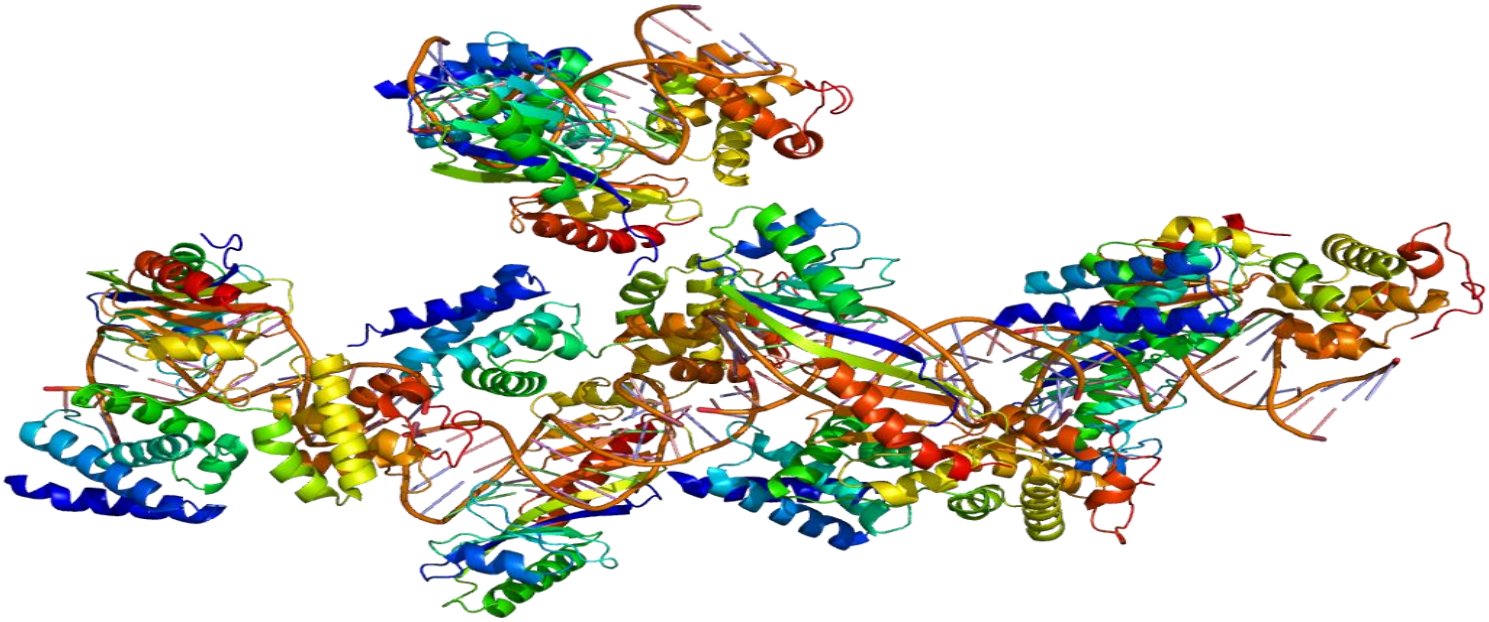
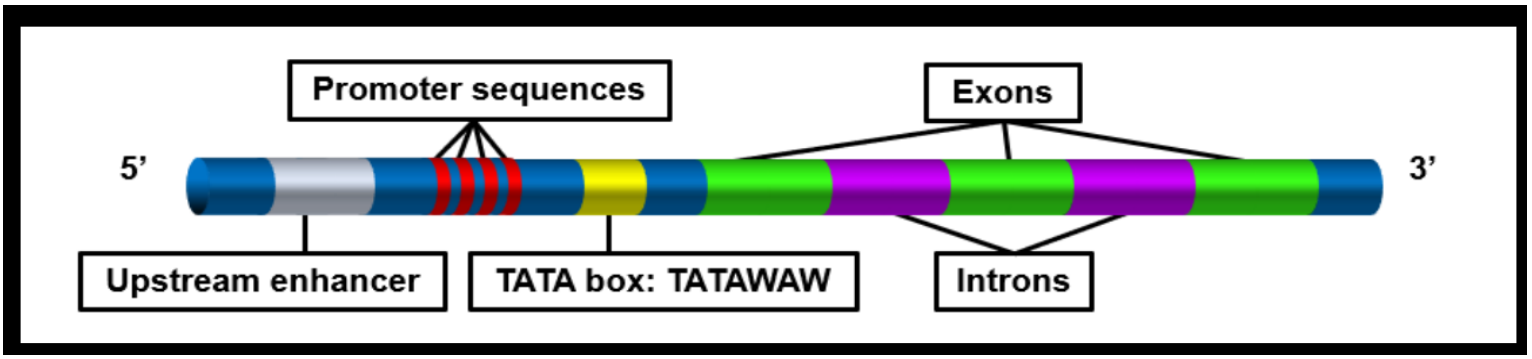


FIGURE: *Structure of the TATA Binding Protein (TBP) or TATA-box binding protein*

The **TATA box** is a specific DNA sequence within the core promoter region, typically located about 25-35 base pairs upstream of the TSS. It has the consensus sequence characterized by repeating T and A base pairs, TATAWAW, where W is either A or T. It plays a crucial role in the initiation of transcription:

i. **Binding Site for TBP**: The TATA box is recognized and bound by the TATA-binding protein (TBP), a subunit of the transcription factor IID (TFIID) complex.

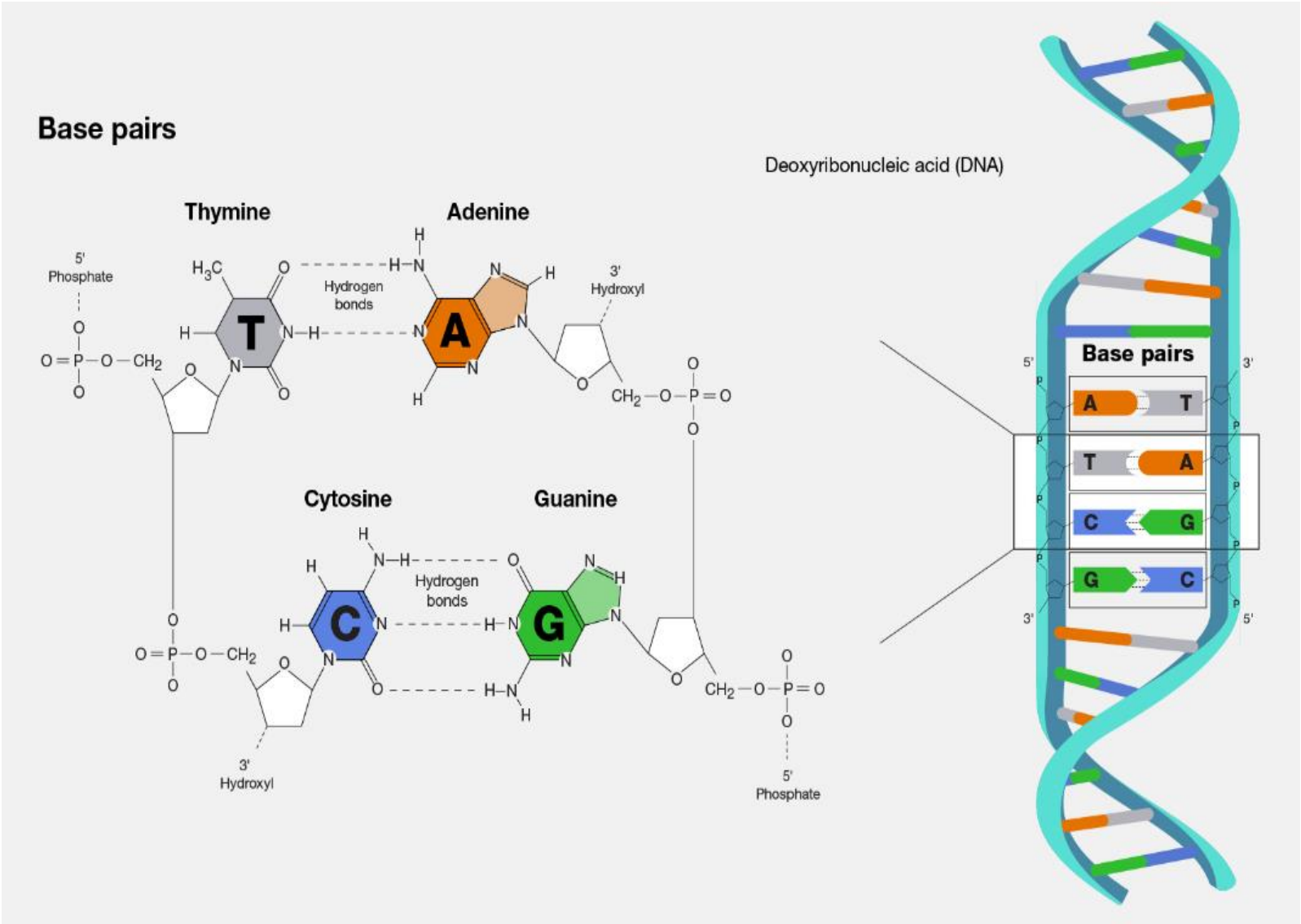
ii. **Recruitment of Transcription Machinery**: The binding of TBP to the TATA box facilitates the recruitment of other transcription factors and RNA polymerase II, forming the pre-initiation complex (PIC).



iii. **Positioning RNA Polymerase II**: The TATA box helps position RNA polymerase II correctly at the TSS, ensuring accurate initiation of transcription.

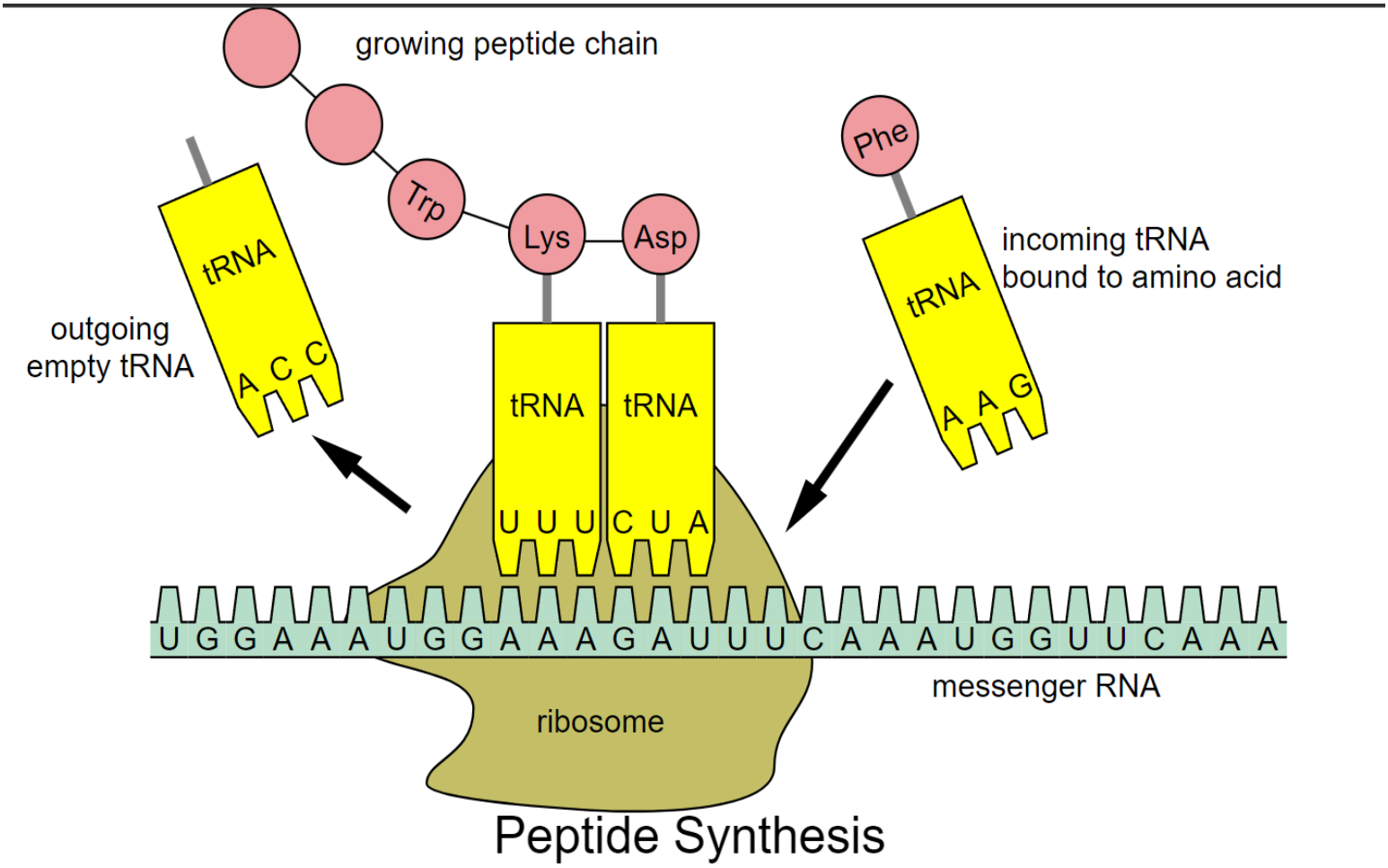
TATA box structural elements. The TATA box consensus sequence is TATAWAW, where W is either A or T.

BEGINNING of REFRESHER COURSE: *A **base pair** (**bp**) is a fundamental unit of double-stranded nucleic acids consisting of two nucleobases bound to each other by hydrogen bonds. They form the building blocks of the DNA double helix and contribute to the folded structure of DNA.*



The chemical structure of DNA base-pairs.

Diagram showing the structure of DNA base-pairs



The interaction of tRNA and mRNA in protein synthesis.

Illustration of Translation or Peptide Synthesis: *A ribosome, a mRNA and lots of tRNAs (each bound to an amino acid), interact to build a peptide (or a protein) molecule.) Note: Transcription is not depicted here.*

END of Refresher Course.

4.3. Interaction and Mutual Influence

The promoter region and the TATA box work together to regulate gene expression:

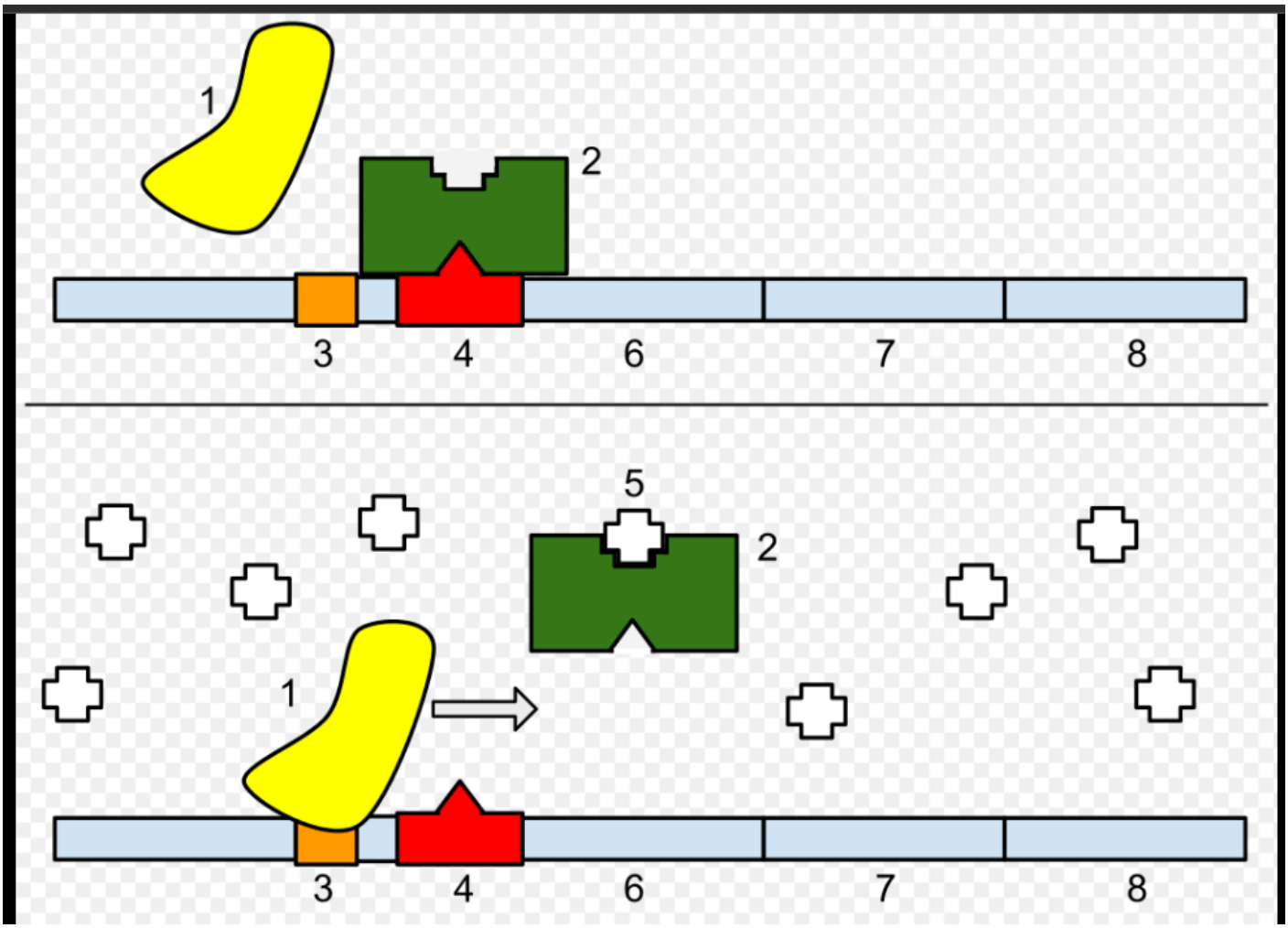
i. **Initiation of Transcription**: The TATA box is a key element in the core promoter that helps initiate transcription by recruiting TBP and other components of the transcription machinery.

ii. **Regulation of Gene Expression**: The presence or absence of a TATA box can influence the responsiveness of a gene to various signals. TATA-containing promoters are often associated with genes that need to be rapidly and dynamically regulated, such as stress-responsive genes.

iii. **Coordination with Other Elements**: The TATA box works in conjunction with other promoter elements (like enhancers and silencers) to fine-tune the transcriptional response. This coordination ensures that genes are expressed at the right time, place, and level.

In summary, the promoter region and the TATA box are essential components of the transcriptional machinery, working together to ensure precise and regulated gene expression.

5. Other promoter elements important in Transcription & gene regulation:



1: RNA Polymerase *2*: Repressor *3*: Promoter *4*: Operator *5*: Lactose *6*: lacZ, *7*: lacY, *8*: lacA.

Top: The transcription of the gene is turned **off**. There is no lactose to inhibit the repressor, so the repressor binds to the operator, which obstructs the RNA polymerase from binding to the promoter and making the mRNA encoding the lactase gene.

Bottom: The gene is turned on. Lactose inhibits the repressor, allowing the RNA polymerase to bind with the promoter and express the genes, which synthesize lactase. Eventually, the lactase will digest all of the lactose, until there is none to bind to the repressor. The repressor will then bind to the operator, stopping the manufacture of lactase.

In addition to the TATA box, several other promoter elements play crucial roles in transcription and gene regulation:

1. **CAAT Box**: Located approximately 75-80 base pairs upstream of the TSS, the CAAT box is recognized by the CAAT-binding transcription factor (CTF) and NF-Y. It enhances the binding of RNA polymerase and other transcription factors, increasing transcription efficiency.
2. **GC-Rich Regions**: These regions, often found in promoters of housekeeping genes, are recognized by the transcription factor SP1. They help maintain a high level of gene expression necessary for basic cellular functions.
3. **Enhancers**: These are distal regulatory elements that can be located far from the promoter. Enhancers increase the transcription of associated genes by looping the DNA to bring the enhancer in contact with the promoter and transcription machinery.
4. **Silencers**: These elements repress gene transcription by binding repressor proteins. Silencers can be located upstream, downstream, or within the gene they regulate.
5. **Insulators**: These elements block the interaction between enhancers and promoters when positioned between them. Insulators help define the boundaries of gene regulatory domains.
6. **TFIIB Recognition Element (BRE)**: Located near the TATA box, the BRE is recognized by TFIIB and helps position RNA polymerase II at the TSS.
7. **Downstream Promoter Element (DPE)**: Found downstream of the TSS, the DPE is recognized by transcription factors and helps initiate transcription in promoters that lack a TATA box.
8. **CpG Islands**: These are regions with a high frequency of CG dinucleotides, often found near the promoters of many genes. CpG islands are typically unmethylated in active genes, allowing transcription to occur.

These elements work together to regulate the initiation and efficiency of transcription, ensuring that genes are expressed at the right time, place, and level.

6. Mutations in the TATA box:

Mutations in the TATA box can significantly impact gene expression by altering the initiation of transcription:

1. **Reduced Transcription Efficiency**: The TATA box is crucial for the binding of the TATA-binding protein (TBP), a component of the transcription factor IID (TFIID) complex. Mutations in the TATA box can

weaken or prevent TBP binding, leading to reduced recruitment of RNA polymerase II and other transcription factors. This results in lower transcription efficiency and decreased gene expression.

2. **Altered Transcription Start Site (TSS)**: Mutations in the TATA box can shift the position of the TSS. This can lead to the production of mRNAs with altered 5' untranslated regions (5' UTRs), which can affect mRNA stability, translation efficiency, and the overall expression of the gene.
3. **Loss of Gene Regulation**: The TATA box is involved in the precise regulation of gene expression. Mutations can disrupt the normal regulatory mechanisms, leading to inappropriate or unregulated expression of the gene. This can have downstream effects on cellular function and contribute to disease.
4. **Context-Dependent Effects**: The impact of TATA box mutations can vary depending on the context of the promoter and the presence of other regulatory elements. For example, some mutations may have a more pronounced effect in the presence of specific enhancers or silencers.
5. **Disease Associations**: Mutations in the TATA box have been linked to various diseases. For instance, certain mutations can lead to reduced expression of tumor suppressor genes, contributing to cancer development.

These effects highlight the critical role of the TATA box in maintaining proper gene expression and how its mutations can lead to significant changes in cellular function and disease.

7. 30 examples of diseases associated with TATA box mutations, along with their characteristics and mechanisms:

Disease	Characteristics	Mechanism
β-Thalassemia	Severe anemia, fatigue, growth retardation	Mutations in the TATA box of the HBB gene reduce transcription efficiency, leading to decreased β-globin production
Hemophilia B Leyden	Blood clotting disorder, excessive bleeding	TATA box mutations in the F9 gene reduce factor IX production, impairing blood clotting
Gastric Cancer	Stomach tumors, weight loss, abdominal pain	TATA box mutations in various genes can lead to dysregulated gene expression, promoting tumor growth

Disease	Characteristics	Mechanism
Spinocerebellar Ataxia	Progressive loss of coordination, balance issues	TATA box mutations in ATXN1 gene affect transcription, leading to neurodegeneration.
Huntington's Disease	Neurodegeneration, movement disorders, cognitive decline	TATA box mutations in HTT gene can alter transcription, contributing to disease progression
Retinitis Pigmentosa	Progressive vision loss	TATA box mutations in genes like RHO affect transcription, leading to retinal degeneration.
Immunosuppression	Increased susceptibility to infections	TATA box mutations in immune-related genes can impair immune response.
Gilbert's Syndrome	Mild jaundice, elevated bilirubin levels	TATA box mutations in UGT1A1 gene reduce enzyme production, affecting bilirubin metabolism.
HIV-1 Susceptibility	Increased risk of HIV infection	TATA box mutations in CCR5 gene can affect receptor expression, influencing HIV entry into cells.
Hepatocellular Carcinoma	Liver cancer, jaundice, abdominal swelling	TATA box mutations in tumor suppressor genes can lead to uncontrolled cell growth.
Oral Cavity Cancer	Tumors in the mouth, difficulty swallowing	TATA box mutations in oncogenes or tumor suppressor genes can dysregulate cell proliferation
Lung Cancer	Persistent cough, chest pain, weight loss	TATA box mutations in genes like EGFR can lead to aberrant gene expression, promoting cancer.
Arterial Hypertension	High blood pressure, risk of heart disease	TATA box mutations in genes regulating blood pressure can affect vascular function.
Myocardial Infarction	Heart attack, chest pain, shortness of breath	TATA box mutations in genes involved in cardiac function can impair heart health.
Thrombophlebitis	Blood clots, vein inflammation	TATA box mutations in coagulation-related genes can increase clotting risk.
Parkinson's Disease	Motor symptoms, tremors,	TATA box mutations in SNCA gene can affect

Disease	Characteristics	Mechanism
	rigidity	transcription, contributing to neurodegeneration.
Alzheimer's Disease	Memory loss, cognitive decline	TATA box mutations in APP gene can alter transcription, affecting amyloid precursor protein production.
Neurofibromatosis Type 1	Tumor formation along nerves	TATA box mutations in NF1 gene can disrupt normal gene expression, leading to tumor growth.
Neurofibromatosis Type 2	Benign tumors on nerves	TATA box mutations in NF2 gene can affect transcription, promoting tumor development.
Melanoma	Skin cancer	TATA box mutations in genes regulating cell cycle can lead to uncontrolled cell proliferation.
Retinoblastoma	Eye cancer in children	TATA box mutations in RB1 gene can disrupt cell cycle control, leading to tumor formation.
Lynch Syndrome	Increased risk of colorectal and other cancers	TATA box mutations in mismatch repair genes can impair DNA repair, increasing cancer risk.
Familial Adenomatous Polyposis	Numerous polyps in the colon, high cancer risk	TATA box mutations in APC gene can affect transcription, promoting polyp formation.
Cystic Fibrosis	Thick mucus production, respiratory issues	TATA box mutations in CFTR gene can disrupt splicing, affecting protein function.
Beta-Thalassemia	Anemia, fatigue	TATA box mutations in HBB gene can disrupt β-globin production.
Sickle Cell Disease	Pain, anemia, organ damage	TATA box mutations in HBB gene can affect hemoglobin production.
Duchenne Muscular Dystrophy	Muscle weakness, progressive muscle degeneration	TATA box mutations in DMD gene can disrupt dystrophin production.
Spinal Muscular Atrophy	Muscle weakness, motor	TATA box mutations in SMN1 gene can affect splicing,

Disease	Characteristics	Mechanism
	neuron degeneration	leading to reduced SMN protein levels.
Marfan Syndrome	Connective tissue disorder, cardiovascular issues	TATA box mutations in FBN1 gene can affect fibrillin production.
Ehlers-Danlos Syndrome	Hypermobile joints, skin elasticity	TATA box mutations in COL5A1 gene can disrupt collagen production.

These examples illustrate the diverse impacts of TATA box mutations on gene expression and disease development.

IMPORTANT QUESTION: Is the function of Transcription factors (TFs) position-dependent to the Transcription Start Site (TSS)?

YES, the function of transcription factors (TFs) is indeed position-dependent relative to the Transcription Start Site (TSS):

Position-Dependent Function

1. **Binding Sites**: The location of TF binding sites relative to the TSS can significantly influence their function. For example, some TFs act as activators when they bind upstream of the TSS but can act as repressors when they bind downstream.
2. **Spatial Arrangement**: The spatial arrangement of TF binding sites can create different regulatory outcomes. This means that even the same TF can have different effects on gene expression depending on its precise position relative to the TSS.
3. **Functional Profiles**: Many TFs, including well-known ones like NRF1, NFY, and Sp1, can either activate or repress transcription initiation based on their position. This positional effect helps guide where and how frequently transcription is initiated.

Examples

- **NRF1**: Typically acts as an activator when bound upstream of the TSS but can repress transcription when bound downstream.
- **YY1**: Has a dual role and can act as both an activator and a repressor depending on its binding position relative to the TSS.

Understanding these position-dependent effects is crucial for decoding the regulatory information in our genome and predicting how mutations in regulatory sequences might impact gene expression and contribute to diseases.

SOURCES:

1.https://bio.libretexts.org/Bookshelves/Microbiology/Microbiology_%28Boundless%29/07%3A_Microbial_Genetics/7.05%3A_RNA_Synthesis_-_Transcription/7.5B%3A_The_Promoter_and_the_Transcription_Machinery

2. *Kornberg RD (August 2007). "The molecular basis of eukaryotic transcription". Proceedings of the National Academy of Sciences of the United States of America.* ***104*** *(32): 12955–61. Bibcode:2007PNAS..10412955K. doi:10.1073/pnas.0704138104. PMC 1941834. PMID 17670940.*

3. Allen BL, Taatjes DJ (2015). "The Mediator complex: a central integrator of transcription". *Nature Reviews. Molecular Cell Biology*. **16** (3): 155–66. doi:10.1038/nrm3951. PMC 4963239. PMID 25693131.

4. Gehring, Alexandra M.; Walker, Julie E.; Santangelo, Thomas J.; Margolin, W. (15 July 2016). "Transcription Regulation in Archaea". *Journal of Bacteriology*. **198** (14): 1906–1917. doi:10.1128/JB.00255-16. PMC 4936096. PMID 27137495.

5. Goodfellow, Sarah J.; Zomerdijk, Joost C. B. M. (2012-06-28), *Basic Mechanisms in RNA Polymerase I Transcription of the Ribosomal RNA Genes*, Subcellular Biochemistry, vol. 61, Dordrecht: Springer Netherlands, pp. 211–236, doi:10.1007/978-94-007-4525-4_10, ISBN 978-94-007-4524-7, PMC 3855190, PMID 23150253, retrieved 2023-10-30

6. Han, Y; Yan, C; Fishbain, S; Ivanov, I; He, Y (2018). "Structural visualization of RNA polymerase III transcription machineries". *Cell Discovery*. **4**: 40. doi:10.1038/s41421-018-0044-z. PMC 6066478. PMID 30083386.

7. https://rna.cd-genomics.com/resource/mapping-of-transcription-start-sites-definition-and-method.html

8. https://www.cell.com/trends/cell-biology/pdf/S0962-8924%2824%2900033-3.pdf

9. Savinkova L, et al., (2013). "An experimental verification of the predicted effects of promoter TATA-box polymorphisms associated with human diseases on interactions between the TATA boxes and TATA-binding protein". PLOSONE. ***8*** *(2):e54626. Bibcode:2013PLoSO...854626S. doi:10.1371/journal.pone.0054626. PMC 3570547. PMID 23424617.*

10. *De Re V, et al., (2017). "P.08.10: Interference of PG2 Tata Box Region with the Serum PG2 Level in Gastric Cancer". Digestive and Liver Disease.* ***49****: e182–e183. doi:10.1016/s1590-8658(17)30534-0. S2CID 79101992.*

11. *Roshan R, Choudhary A, Bhambri A, Bakshi B, Ghosh T, Pillai B (August 2017). "microRNA dysregulation in polyglutamine toxicity of TATA-box binding protein is mediated through STAT1 in mouse neuronal cells". Journal of Neuroinflammation.* ***14*** *(1): 155. doi:10.1186/s12974-017-0925-3. PMC 5543588. PMID 28774347.*

12. Wu C, Liu Z, Ma L, Pei C, Qin L, Gao N, Li J, Yin Y (August 2017). "MiRNAs regulate oxidative stress related genes via binding to the 3' UTR and TATA-box regions: a new hypothesis for cataract pathogenesis". BMC Ophthalmology. ***17*** *(1): 142. doi:10.1186/s12886-017-0537-9. PMC 5556341*

13. : Bae S-H, Han HW, Moon J (2015) Functional Analysis of the Molecular Interactions of TATA Box-Containing Genes and Essential Genes. PLoS ONE 10(3): e0120848. doi:10.1371/journal. pone.0120848.

14. Sharan R (4 January 2007). "Analysis of Biological Networks: Transcriptional Networks – Promoter Sequence Analysis" (PDF). Tel Aviv University.

15. Vaishnav ED, et al. (March 2022). "The evolution, evolvability and engineering of gene regulatory DNA". *Nature*. **603** (7901): 455–463. Bibcode:2022Natur.603..455V. doi:10.1038/s41586-022-04506-6. PMC 8934302.

16. L. K. Savinkova, E. B. Sharypova, N. A. Kolchanov. On the Role of TATA Boxes and TATA-Binding Protein in Arabidopsis thaliana. Published: 22 February 2023. *Plants* 2023, *12*(5),1000. This article belongs to the Section Plant Molecular Biology.

17. https://bmcmedgenet.biomedcentral.com/articles/10.1186/s12881-020-01106-x.

18. Antonarakis SE, Irkin SH, Cheng TC, Scott AF, Sexton JP, Trusko SP, Charache S, Kazazian HH (1984). "beta-Thalassemia in American Blacks: novel mutations in the "TATA" box and an acceptor splice site". *Proceedings of the National Academy of Sciences of the United States of America*. **81** (4): 1154–8. Bibcode:1984PNAS...81.1154A. doi:10.1073/pnas.81.4.1154. PMC 344784. PMID 6583702.

19. De Re V, Magris R, et al., (2017). "P.08.10: Interference of PG2 Tata Box Region with the Serum PG2 Level in Gastric Cancer". *Digestive and Liver Disease*. **49**: e182–e183. doi:10.1016/s1590-8658(17)30534-0. S2CID 79101992.

20.https://www.academia.edu/11198313/The_Mechanism_by_which_TATA_Box_Polymorphisms_Associated_with_Human_Hereditary_Diseases_Influence_Interactions_with_the_TATA_Binding_Protein.

21. https://www.nature.com/articles/s41586-024-07662-z.pdf. JULY 2024.

22. https://www.nature.com/articles/s41576-024-00769-z. AUGUST 2024.

23. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE199431. APRIL 2024.