Human gene JAK2 is a promising model for studying regulation of genes expression by a mechanism of wave collapse on the DNA.

Denis Semyonov.
E-mail address: dasem@mail.ru

In human gene JAK2 the fragment with regulatory structure GTATGTGT is repeated more than 10 times. The mutation V617F (G→T) in this site belonging to an 14 exon, is associated with so called myeloproliferative diseases. E.g., if the mutation leads to violation of the purine-pyrimidine alternation, the respective site loses the ability of the rearrangement and becomes non-regulatory. Thus, contribution of the mutation to oncogenesis can relate to a disturbance of the mechanism of regulation of the gene expression.

In search of an object where genes expression is regulated via a mechanism of rearrangements in the sites with alternated purines and pyrimidines, I have found gene JAK2. In this gene the same fragment GTATGTGT is repeated more than 10 times. This fragment represents a site with alternated purines and pyrimidines. As I have suggested in my previous work [1], at such DNA sites it is possible collapse of torsional waves, co-operative rearrangement of the whole fragment to Hoogsteen base pair, and it is a probability of the subsequent transformation of the fragment into Z-DNA.

Although the gene itself is rather long (about 150,000 base pairs), random appearance of so large number of the above fragments is almost improbable.

The above sites are distributed over the whole gene, which suggests unspecific character of their putative binding with a regulatory molecule, which can serve as an obstacle for transcription process and thereby decrease the genes expression rate. Besides exact copies of GTATGTGT fragment, the gene contains approximately 100 copies of this fragment containing various mutations, which probably exits as a trash, i.e. fragments lacking regulatory activity. E.g., if the mutation leads to violation of the purine-pyrimidine alternation, the respective site loses the ability of the rearrangement and becomes non-regulatory; such sites are not rare in the JAK2 gene. As well, rather frequent are mutant copies where purine-pyrimidine alternation is conserved; possibly, these sites retain the ability of the rearrangement and can be recognized by a regulatory molecule. Thus, it is possible that some mutant copies of the regulatory site participate in regulation of the transcription rate.

Interestingly, one of the mentioned sites is within an exon fragment GTATGTGTC coding for tripeptide Val-Cys-Val. It is commonly accepted that regulatory elements very rarely present in exons. Nevertheless, let us assume the discussed site as regulatory, and consider mutations in this site as those violating regulatory functions in DNA but not simple amino acid residue replacement. The more

remarkable thing is that the mutation in this site belonging to an exon, is associated with so called myeloproliferative diseases. It is believed that mutation G→T replacing GTATGTGTC to GTATGTTTC (Val→Phe) results in an increase of the JAK2 signal leading to redundant proliferation of cell growth various type.

If a sequence has a feature of a regulatory site and is present in the gene in many copies, it is reasonable to suggest that the sequence is important for the regulation of gene JAK2 expression. Such kind DNA sites are expected to bind regulatory proteins. Regulatory sites stretched along the gene and regulatory proteins bound at these sites could make obstacles for transcription and thereby regulate the gene expression. There assumption is confirmed by a finding, which has shown that the fragment GTATGTGT under discussion (see above) is a component of binding sites for at least four different transcription factors, namely, HIF1a:ARNT, BATF:JUN, FOSL2, JUN (var.2). [2]. All mentioned transcription factors are regulated by proteins from the STAT family and are, in turn, controlled by JAK2. One can assume that GTATGTGT sequences in the JAK2 gene have affinity to the above transcription factors. This can provide a feedback regulation of the gene so that the level of its expression should decrease upon the activation of the transcription factors.

Thus, contribution of the mutation to oncogenesis can relate to a disturbance of the mechanism of regulation of the gene expression. In particular, the mutation can lead to a disturbance of the regulatory site, which, in turn, can lead to the enhancement of expression. The latter is the case at myeloproliferative diseases associated with the mutation V617F (G→T) in the JAK2 gene.

Mutation G→T probably originates from oxidative lesions of guanines occurring nonenzymatically. When guanine becomes oxidized to 8-oxoguanine, it can form a base pair with adenine. The resulting pair 8-oxoG-A is missed by DNA repair enzymes [3] and oxidized G in DNA eventually becomes replaced with T.

The Hoogsteen base pair formation can facilitate replacement G→T since orientation of G (and of 8-oxoG) in Hoogsteen base pair 8-oxoGC is favorable for the following 8-oxoG base pairing with A. At least, this mechanism can be supported by the respective structural formula (Fig. 1). Base pair 8-oxoGC presented in the Figure is formed in a geometry similar to that of Hoogsteen base pair, and the Figure shows that such pair can be very stable (due to a possibility of the formation of three hydrogen bonds) and live up to the replication. If during replication 8oxoG will not rearrange, it will be base paired with A, and it will finally result in the replacement G→T.

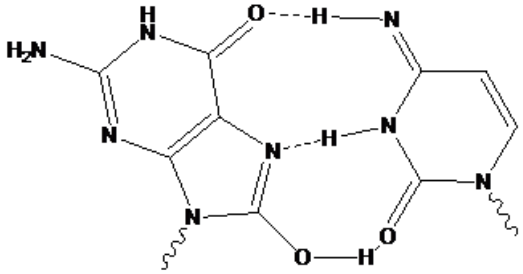

Figure 1. Hypothetical pair 8-oxoGC Hoogsteen.

Protonation of cytosines should facilitate the GC Hoogsteen base pairs formation upon decrease of pH [4]. Comparison between the duplexes GTATGTGT\ACACATAC and GCGCGCGC\GCGCGCGC shows that in the latter one protonation of the neighboring cytosines in unlikely since they are too close to each other in contrast to the former duplex. Regulatory sites in the JAK2 gene can be pH-sensitive. One can suggest the following chain of events:
1. Protonation of cytosines at a DNA site complementsry to GTATGTGT;
2. Turn of the base pairs at this site to Hoogsteen ones;
3. Binding of a regulatory protein to GTATGTGT;
4. Inhibition of RNA polymerase at this site with the bound regulatory protein.
For example, regulatory protein HIF1a:ARNT is activated under hypoxia conditions. Hypoxia leads to the occurrence of glycolysis, which, in turn, results in decrease of pH. The effect of pH on the JAK2 expression might contribute to the regulation of the level of the response to hypoxia.

Acknowledgements:
The author is grateful to D.M. Graifer for helpful discussion and his help in the text preparation.

References:

1. Semyonov D. Collapse of the Waves on “soft” DNA Sites Can be a Physical Basis of the Regulation of Genes Expression. viXra:1704.0155
2. Buroker N.D. Computational STAT4 rSNP analysis, transcriptional factor binding sites and disease. Journal of bioinformatics and diabetes. 2016. Vol-1 Issue – 2 p. 18-53.
3. D.O. Zharkov, A.P. Grollman. Repair of Oxidative DNA Damage. in Encyclopedia of Genetics, 2001
4. E.N. Nikolova, G.B. Goh, C.L. Brooks III, and H.M. Al-Hashimi. Characterizing the Protonation State of Cytosine in Transient G•C Hoogsteen Base Pairs in Duplex DNA. J Am Chem Soc. 2013 May 8; 135(18): 6766–6769