coopern/Shutterstock.com

Meier-Stephenson's work could provide the basis for novel cancer treatments and antivirals.

G4-quadruplexes

How tangled knots of DNA help regulate and protect our genes

- One of the ways in which DNA regulates and protects itself is by forming tangled knots called G4-quadruplexes.
- The exact function of these regions, and the proteins which interact with them, is still unknown.
- Dr Vanessa Meier-Stephenson at the University of Alberta, Canada, is working to better understand the various functions of G4-quadruplexes and their binding proteins.
- This knowledge could help scientists develop drugs targeting cancer or chronic viruses.

DNA encodes all the information needed to make every protein and enzyme in a living organism, acting as instructions for the basic building blocks in our body, including genes. But many of these genes are not needed all the time. So, our genetic machinery has developed the ability to effectively 'switch' different coding regions 'on' or 'off' – to either start or halt the expression of a gene. A process called transcription is the first stage of gene expression. DNA can regulate its transcription in lots of different ways, such as coiling around structural proteins or interacting with enzymes.

Now, research by Dr Vanessa Meier-Stephenson, an Assistant Professor at the University of Alberta, Canada, examines novel ways to regulate some of these pathways. This involves DNA folding itself into intricate 3D structures called G4-quadruplexes. Meier-Stephenson's findings could result in the development of new targeted treatments for conditions like cancer or chronic viral infections.

Tangles in the helix

Usually, the DNA molecule is a long sequence of the nucleotides (single units of the DNA code) adenosine, tyrosine, cytosine, and guanosine twisted into a perfect helix. However, in areas where the chain contains an abundance of the nucleotide guanosine, or G, those sections can fold the DNA molecule into a three-dimensional structure known as a G4-quadruplex (G4Q). The G4Q complex forms when four guanosines arrange themselves in sheets stacked on top of each other in various orientations. This causes the DNA strand to form bends and loops. The stability of the G4Q complex is affected by loop length, where longer loops tend to make the complexes more unstable. The shape and orientation of these loops influences which proteins and ions can interact with them. G4Qs have been associated with important regulatory regions, which is what makes them so interesting to genetics researchers.

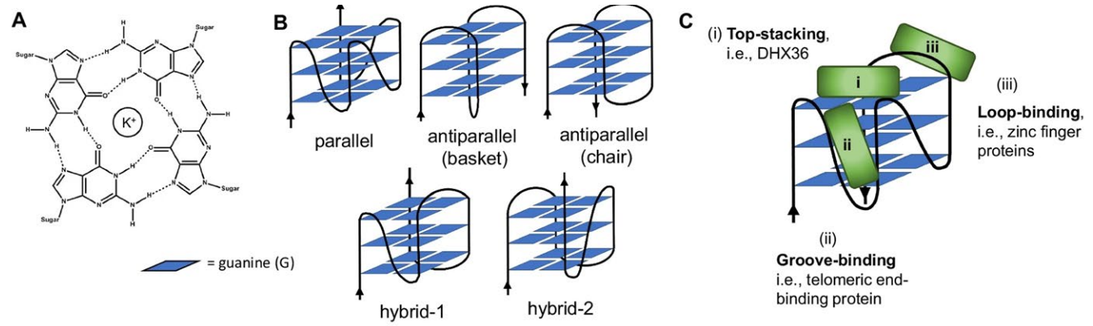
The formation of G4Qs is not fully understood, but previous studies have demonstrated that their stabilisation and destabilisation in vitro can be influenced by temperature fluctuations. As these studies were done under artificial conditions outside a living cell, we don't know for certain whether they reflect the natural G4Q development process. After all, there are many things within a living cell that can also interact with DNA, like the proteins which bind to and copy or translate the DNA code. Also, there could be other G4Q shapes that we have no way of seeing or measuring in the laboratory. Nevertheless, once formed, G4Qs are stable as a solitary DNA double helix.

Structure, function, and binding proteins

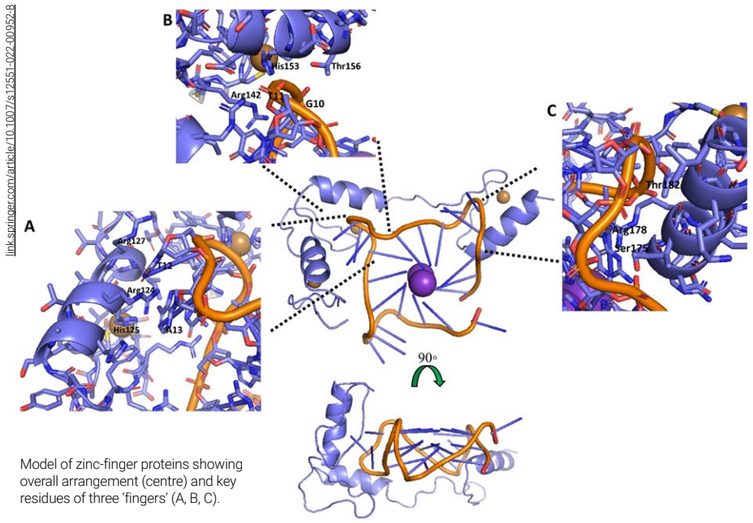
Over 700,000 sequences have been discovered within the human genome that are involved in regulating our DNA. They are found in regions of the DNA sequence requiring more intense regulation, such as a gene promotor. A promotor is a region of DNA at the beginning of a gene where the transcription proteins bind to begin expressing the gene. By manipulating this region of the DNA sequence, a regulatory mechanism can prevent proteins from binding to the DNA – and therefore 'switch' the gene off.

G4Qs are as stable as a solitary DNA double helix and have been associated with important regulatory regions.

G4Q complexes have been shown to form 'knots' in the DNA, which change the way proteins interact with it or protect the ends of the DNA molecule from degradation. For example, they help stabilise the telomeres. Telomers are regions at the end of the DNA strand that protect the DNA helix's edges. Without the telomeres, proteins in the cell



Schematic of G4-quadruplexes: these structures interact with various cellular proteins, which bind in a number of different ways, including top-stacking, groove-binding, and loop-binding.



Model of zinc-finger proteins showing overall arrangement (centre) and key residues of three 'fingers' (A, B, C).

Exploring DNA regulatory mechanisms helps build on our knowledge of how DNA switches genes on and off.

would see the DNA as a broken strand in need of recycling. The G4 complexes are also known to attract and anchor the proteins required for gene expression. The variety of G4Q-associated sites, and their importance, suggest that they are involved in the selective regulation of critical genetic functions. The proteins attracted to them may carry out specialised regulatory tasks, such as 'turning on' a gene or recruiting another important factor.

Known G4-quadruplex binding proteins

The specialised proteins associated with G4Qs, known as G4BPs, can be identified either by their stabilising or destabilising influences on the G4Q complex or by the region to which they bind. Since many of these new proteins' functions are often unknown, designation via associated gene region is more common. For example, POT1 is a protein associated with the telomere G4Q sequences, and it stabilises the G4Q to prevent telomere degradation. Replication Protein A is another G4BP heavily involved in DNA replication, repair, and recombination.

Meier-Stephenson's group is interested in exploring several proteins and protein complexes associated with G4Qs, especially a class of G4BPs known as zinc-finger proteins. These proteins appear to have an 'anchoring' function, and several are associated with promotor region G4Qs. Despite their integral roles in many cellular processes, there are limited models and structural data on these protein-G4Q complexes. Her research group is building upon prior work to increase knowledge around these important complexes. Specifically, by analysing and modelling the structure of the protein-G4Q complex at the molecular level, the group is establishing the exact factors contributing to the complex's stability, what components of the interaction make it specific for a certain G4Q and how these structural insights can be used to create therapies.

Potential drug targets

Exploring DNA regulatory mechanisms helps build on our knowledge of how DNA switches genes on and off. Understanding these mechanisms could inform future drug targets, for example, by disabling a protein known to switch on cancer-associated genes or other disease-causing processes.

G4Qs and their associated binding proteins are critical to many cellular processes. Their varied binding orientations confer the selectivity and versatility required of these complex interactions. Studying this specific regulation and control of important processes, such as turning genes 'on' or 'off' in response to the environment, could inform future genetics research and potential drug targets. Significantly, Meier-Stephenson's work could provide the basis for novel treatments for cancer, disease, or chronic viral infections.

Personal response

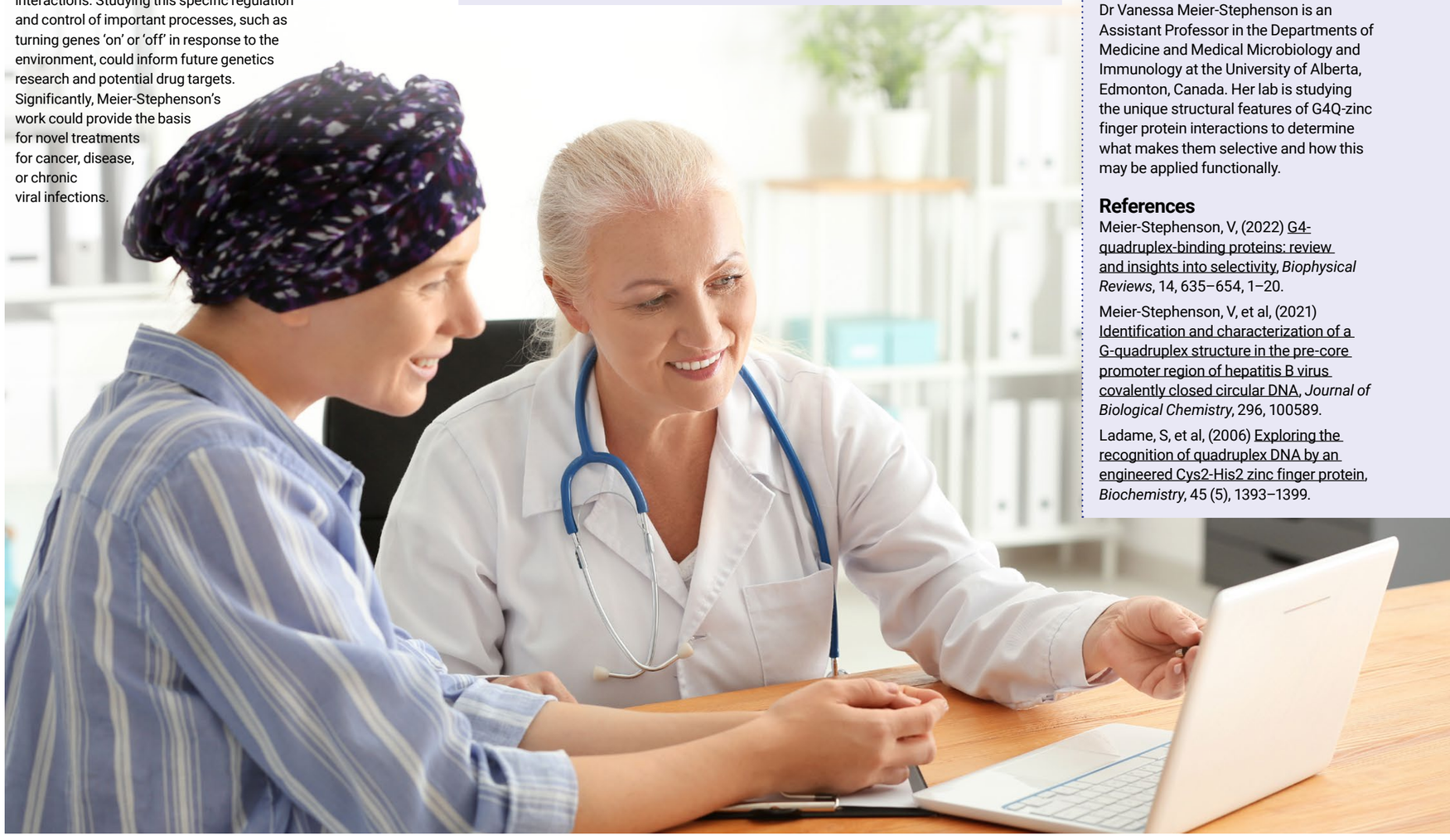
What further studies are needed to understand the newly discovered G4-quadruplexes?

I guess they're not technically new, just newly recognised as such. More and more studies are linking the G4Q's presence with a functional outcome. With the 700,000 putative sequences, pulling out those of greatest interest to focus on is a strategic decision. But interestingly, some older data can be revisited and potentially explained by some of these new insights.

Current drug discovery approaches in the anti-cancer realm looks at small-molecule binders with an appropriate downstream assay, then can start to modify components/ligands to improve binding and/or selectivity. It's an important approach to get moving on finding a target and focusing energies on the endpoint goals. The problem is, if you don't get a hit, you're not necessarily further along in understanding why or how those G4Qs are playing a role. We're trying to focus some energy on what happens with the 'natural binders' to see if we can figure out what it would take to interfere with it.

Could targeted cancer drug therapies based on G4BPs and their binding sites have fewer side effects than current treatments?

It's possible, but as it stands, many of these small molecules have lots of off-target effects. By understanding how the natural protein binder does this *selectively*, we can also learn about what it takes to design binders that are specific to certain targets.



Details

e: meierste@ualberta.ca
t: +1 780 492 8599
w: sites.google.com/view/meierstephensonlab

Funding

- University of Alberta
- Li Ka Shing Institute of Virology

Collaborators

- Dr Trushar Patel, University of Lethbridge
- Dr Lorne Tyrrell, University of Alberta

Bio

Dr Vanessa Meier-Stephenson is an Assistant Professor in the Departments of Medicine and Medical Microbiology and Immunology at the University of Alberta, Edmonton, Canada. Her lab is studying the unique structural features of G4Q-zinc finger protein interactions to determine what makes them selective and how this may be applied functionally.

References

Meier-Stephenson, V, (2022) G4-quadruplex-binding proteins: review and insights into selectivity, *Biophysical Reviews*, 14, 635–654, 1–20.

Meier-Stephenson, V, et al, (2021) Identification and characterization of a G-quadruplex structure in the pre-core promoter region of hepatitis B virus covalently closed circular DNA, *Journal of Biological Chemistry*, 296, 100589.

Ladame, S, et al, (2006) Exploring the recognition of quadruplex DNA by an engineered Cys2-His2 zinc finger protein, *Biochemistry*, 45 (5), 1393–1399.