Explor Anim Med Res,
Vol.10, Issue - 1, 2020, p. 18-23

ISSN 2277- 470X (Print), ISSN 2319-247X (Online)
Website: www.animalmedicalresearch.org

Review Article

I-MOTIF DNA: SIGNIFICANCE AND FUTURE PROSPECTIVE

Sushmita Nautiyal[1], Vishal Rai[2*], Sudipta Bhat[3], Ravi Kumar[4], Mudasir Mohd Rather[5], M.Sankar[6]

Received 25 November 2019, revised 19 June 2020

ABSTRACT: Since the first report of its existence emerged in the early 1990s, i-motif DNA has gained much interest among the other non-canonical DNA structures in recent years. It is a quadruplex structure made up of two parallel duplexes formed by folding of cytosine-rich sequences of the nucleic acid with the cytosine•cytosine+ base pair as its building block. Earlier known to exist only at acidic pH, genomic sequences that form stable i-motif structures even at neutral pH under certain conditions have been reported recently. It is visible during G1 stage of cell cycle and is also seen in promoter sites and in telomeres. It is the only known nucleic acid structure involving systematic base intercalation. Owing to its sharp, fast and reversible pH-driven conformational changes it has already inspired the designs of various nanomachines and has got wide applicability in fields of nanotechnology and analytical chemistry. Its recent confirmation *in vivo* has given insight to its potential role in various important biological processes like replication, regulation of oncogene expression, and telomere functions. Therefore, the targeting of i-motif DNA and active search for appropriate ligands interacting with this non-canonical structure is an emerging area of research in medicinal and nucleic acid chemistry.

Key words: I-motif, Ligands, Nanomachines, Oncogene, Quadruplex, Telomere.

INTRODUCTION

Since 1953, when James Watson and Francis Crick famously uncovered the iconic double helix structure of DNA, it has captured the imagination of all humankind. B-DNA form occurs most frequently under physiological conditions. But now it is well known that short stretches of DNA can exist in other shapes as well. Alternate non-B form conformations including G quadruplex (G4) and intercalated motif (i-motif) structures have been reported *in vivo*, i-motif being the recent one to be detected in the nuclei of human cell (Zeraati *et al*. 2018). In genomic DNA, guanine rich sequences fold into G quadruplex (Henderson *et al*. 1987) and their complementary cytosine rich sequences also form quadruplex structures known as i-motifs, also known as i-tetraplex, intercalated motif, i-DNA or twisted knot.

These short, recurring patterns in DNA might play some important role in various biological functions. I-motif structure is formed from the folding of cytosine rich nucleic acid sequences and the core structure consists of two parallel duplexes intercalated in an antiparallel manner. It is the only known DNA structure consisting of parallel-stranded duplexes that are held together through intercalated base pairs, unlike other structures such as B-DNA or G-quadruplex, which are spatial arrangements of DNA strands held together by means of stacked base pairs. The building block of i-motif DNA is the base-pair involving one neutral (deprotonated) cytosine and one positively charged (protonated) cytosine at N3 position resulting in C·C+ base pair, which is stable because of the formation of three hydrogen bonds which allows the formation of parallel duplexes. Till now the biological relevance of I-motif has been questioned due to its dependence on acidic conditions *in vitro*, but now it is confirmed that it can be formed at physiological pH under certain conditions of molecular crowding and negative superhelicity induced during transcription (Zeraati *et al*. 2018).

[1]Division of Veterinary Virology, ICAR-Indian Veterinary Research Institute, Mukteswar, Uttarakhand, India. [2]Division of Biological Products, ICAR-Indian Veterinary Research Institute, Bareilly, Uttar Pradesh, India. [3]Division of Biological Standardization, ICAR-Indian Veterinary Research Institute, Bareilly, Uttar Pradesh, India. [4,5]M.V.Sc., ICAR-Indian Veterinary Research Institute, Bareilly, Uttar Pradesh, India. [6]Division of Parasitology, ICAR- Indian Veterinary Research Institute, Bareilly, Uttar Pradesh, India.
[*]Corresponding author, e-mail: roger.vish@gmail.com

I-MOTIF DNA: STRUCTURAL FEATURES

Gehring *et al*. (1993) characterised the first DNA i-motif for the hexameric sequence d(TCCCCC) forming an intercalated quadruple-helical tetramolecular structure under acidic conditions, consisting of two parallel-stranded duplexes intercalated in an antiparallel orientation and held together by hemi-protonated cytosine-cytosine+ (C:C+) base pairs (Fig. 1). Till date, a number of i-motif structures has been determined by crystallographic and NMR methods. I-motifs may be intermolecular formed from the association of two (dimers) or four (tetramers) separate DNA strands or may be intramolecular (monomer) due to the spatial arrangement of four different C-tracts within the same strand. Unlike B-DNA in which the distance between consecutive base pairs is 3.4 Å and right-handed helical twist angle is 36°, the i-motif DNA has 3.1 Å distance between the consecutive base pairs and right-handed helical twist angle of~12–20° (Berger *et al*. 1996).

As with other nucleic acid structures, stability of i-motif structure depend on many factors, including sequence nature, temperature, and ionic strength of the environment. Key interactions for i-motif stability are the hemi-protonated C:C+ base pairs. A high stability is conferred by the three hydrogen bonds of the C:C+ base pair. Base-pairing energy (BPE) for the C:C+ base pair is 169.7 kJ/mol, which is higher than both the BPEs of canonical Watson-Crick G·C (96.6 kJ/mol) and neutral C·C (68.0 kJ/mol) (Yang and Rodgers, 2014). I-motifs are not affected by the nature of the cation (as G quadruplex) but they are affected by the ionic strength of the solution. An increase in the concentration of NaCl from 0 to 100 mM at a pH close to the pKa of cytosine destabilizes i-motif structure (Mergny *et al*. 1995), but higher NaCl concentrations (300 mM) do not cause further destabilization (Mergny *et al*. 1995, Day *et al*. 2014).

BIOLOGICAL RELEVANCE

Location across the genome

I-motifs are found to be frequently clustered in or near regulatory regions of numerous genes & telomeres, indicating involvement in a variety of genome functions. Studies have shown that sequences having higher potential for forming i-motif structures are not randomly located; instead, they are particularly enriched in the promoters of certain genes (especially in the promoter region of oncogenes), suggesting their possible role on certain regulatory mechanisms of gene expression (Wright *et al*. 2017, Fleming *et al*. 2017). However, these studies are not comprehensive, and a better understanding is still necessary to achieve a more accurate mapping of i-motif occurrence along the genome. It has been reported that i-motif structures formed within the promoter regions of *c-Myc* gene and *Bcl-2* gene respectively, could modulate their gene expression (Dexheimer *et al*. 2009, Kendrick *et al*. 2014).

Existence of i-motif structures *in vivo*

Due to the lack of experimental evidence and their dependence on acidic pH, the relevance of i-motif structures was largely questioned till now. Very recently, an antibody fragment (iMab) that recognizes i-motif structures with high selectivity and affinity has been generated and characterized by Christ's and Dinger's group, enabling the detection of i-motifs in the nuclei of human cells. They also demonstrated that the in vivo formation of i-motif structures is cell-cycle and pH dependent. Using iMab for immunofluorescent staining in three different cell lines and evaluating the number of foci in cells arrested in three different stages of the cell cycle (early S phase and G0/G1 andG1/S boundaries), the authors suggested the association of i-motif formation with transcription and the resolving of i-motif structures during DNA replication/S-phase (Zeraati *et al.* 2018). Earlier studies for detection of number of G4 foci in cells by immunofluorescence assays with BG4 antibody showed that their number was higher during S-phase, in which the major event is DNA replication (Biffi *et al.* 2013). This suggested different occurrence of i motif and G4 structures during cell cycle indicating that they might have opposite roles in regulation of gene expression (Zeraati *et al*. 2018, Cui *et al*. 2016, Sutherland *et al*. 2016).

I-motif interaction with ligands and protein

Either due to the low stability of i-motif at physiological conditions or due to its very compact structure (where planar ligands which usually stack on base pairs cannot be introduced easily), there are very few well-documented examples of specific i-motif binding ligands. Carboxyl-modified single-walled carbon nanotubes (CSWNTs) are known to be the first selective i-motif ligand. The thermal stability of the intramolecular i-motif formed by the C-rich human telomeric sequence at acidic pH is remarkably increased by the presence of CSWNTs. They also inhibit duplex formation between complementary human telomeric C and G-rich sequences and induces the formation of this i-motif structure at pH 8.0 (the favourable electrostatic interactions between the C:C+ base pairs and the substantial decrease of the pKa of the C:C+ base pairs by the negatively charged

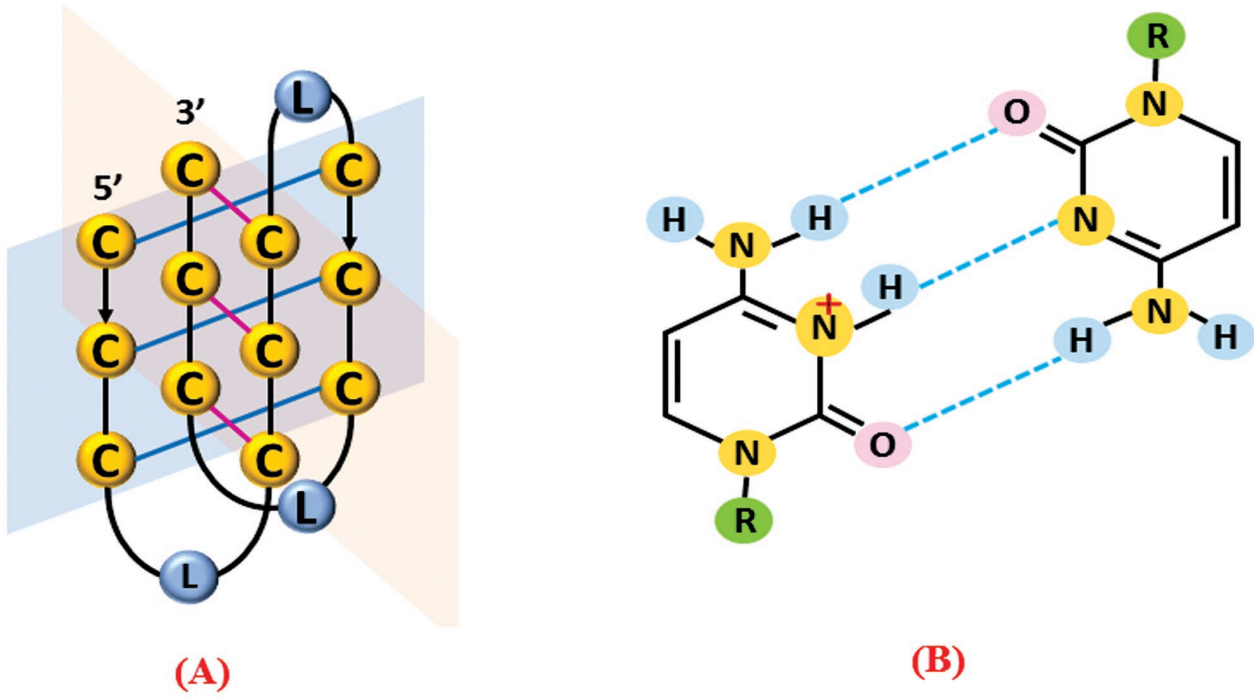

Fig. 1. Schematic diagrams representing: (A) Intramolecular i-motif, where C and L represent the cytosine residues and loop region with any DNA base, respectively. (B) A base pair with hemiprotonated cytosine–cytosine+.

CSWNTs explains their stabilization effect) (Assi *et al*. 2018). A study by Chen *et al.* (2012) showed the inhibition of telomerase activity and interference with telomere functions by the i-motif structures in the presence of CSWNTs leading to senescence and apoptosis in cancer cells *in vitro* and *in vivo*. Another study led to the identification of a compound IMC 48 after screening a library of 1990 compounds, which binds and stabilizes the i-motif structure formed by the C-rich sequence of BCL2 gene promoter (Kendrick *et al*. 2014), while a similar compound (IMC 76) was found to favour an alternative hairpin conformation formed by the same sequence. The dynamic equilibrium between these two structures can be shifted by these molecules which have opposite effects; IMC-48 leads to the activation of gene expression whereas IMC-76 markedly suppresses the levels of BCL2 mRNA (Kendrick *et al*. 2014). Through a pull-down assay aimed to find proteins involved in transcription and showing affinity for the i-motif structure formed by the BCL2 promoter oncogene, the BCL2 activating transcription factor heterogeneous nuclear ribonucleoprotein LL (hnRNP LL, a pre-mRNA splicing factor capable of binding and stabilizing BCL2 mRNA) was identified which is one of the very few i-motif binding protein studied in depth. It was found to bind exclusively to the i-motif structure, neither to the BCL2 promoter forming a duplex nor to mutated single strand DNA unable to fold into an i-motif (Kang *et al.* 2014).

Interference with telomerase activity

Telomerase is a ribonucleoprotein that is responsible for the addition of a species-dependent telomere repeat sequence to the 3' end of telomeres and maintenance of genomic integrity in normal cells. Telomeres progressively get shortened during successive cell divisions thereby inducing chromosomal instability. Telomerase activity is higher in most of the cancer types as well as in cancer stem or stem like cells as compared to normal human cells (including stem cells), which maintain telomere at longer lengths than the cancer cells. These features make telomerase an effective target for anticancer therapeutics (Jafri *et al*. 2016). It has been proposed that the formation of i-motifs on one strand is promoted by the formation of one or more G-quadruplex structures at the 3'-terminal G-rich region of the other strand. Their formation at the end of telomeres leads to telomerase inhibition (Fig. 2) (Amato *et al*. 2014). While some studies have shown that the stabilization of certain human telomeric G4 topologies (tandem repeats of the unit 5'-ATTGGG-3'/3'-TAACCC-5') with ligands might lead to the inhibition of telomerase activity; investigation of the effect of targeting the complementary C-rich strand

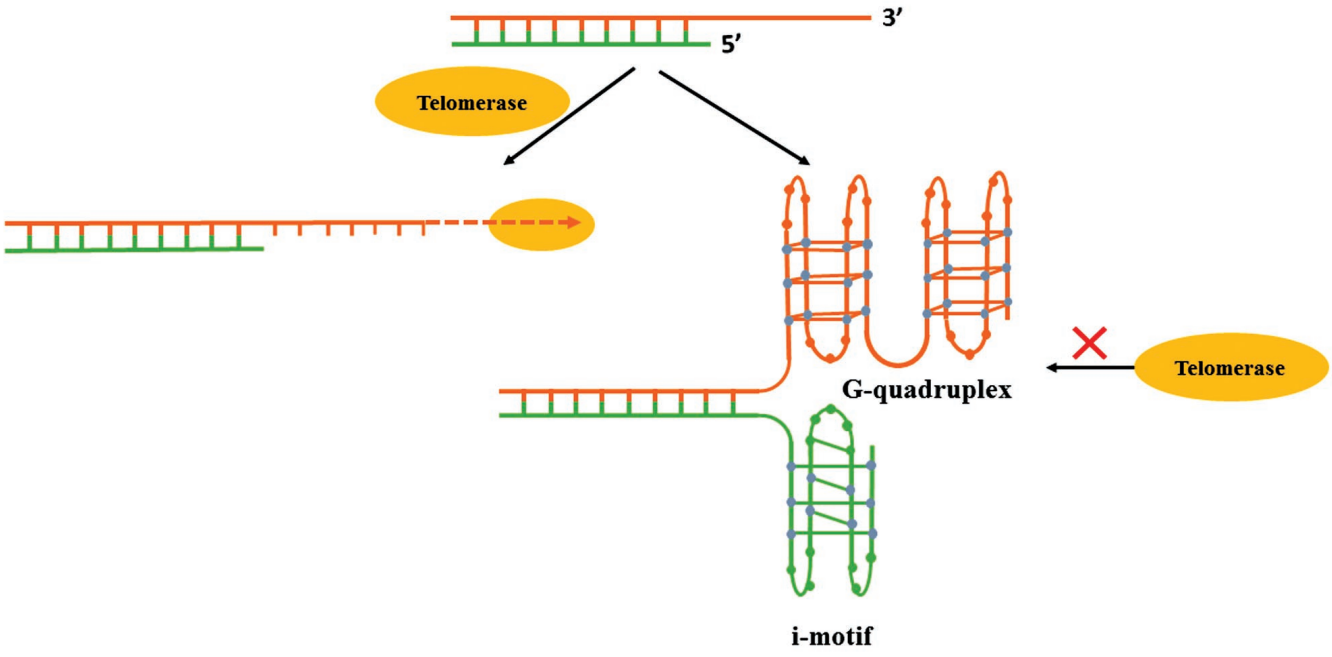

Fig. 2. Mechanism of telomerase inhibition due to G-quadruplex and i-motif structure at the end of telomeres. Concept adapted from Amato *et al*. 2014.

has not been done in depth. Thus as mentioned earlier, the study conducted by Chen *et al*. in 2012 was the first to investigate telomerase activity on i-motifs formed by C-rich human telomeric sequences stabilized by CSWNT. Various observations also suggest possible role of i-motif DNA in directing centromere location and in providing particular architectural features to the centromere (Henikoff *et al*. 2017, Assi *et al.* 2018).

Role of i-motif in transcription

Recent findings by Zeraati *et al.* (2018) showed the highest level of i-motif in late G1 phase, characterized by high levels of transcription and cellular growth. This observation correlates with other findings describing the regulatory role of i-motif structures at the promoters of several proto-oncogenes (Sutherland *et al*. 2016, Roy *et al.* 2016). It supports the concept that i-motifs may act as scaffolds for the binding of transcription factors during transcription (Zeraati *et al*. 2018). Their study also showed resolution of i-motif structures during S phase indicating mutual exclusivity of G4 and i-motif structures as G4 occurs predominately during the S phase.

Effect on DNA biosynthesis

There are studies which suggest that i-motifs could modulate DNA replication *in vivo* and pose a greater impact compared to other secondary structures such as hairpins or mixed-type G4s. When different i-motif-forming sequences were inserted in the template strand of the replication reaction, they were found to stall DNA polymerase and thus impede DNA replication or repair (Takahashi *et al*. 2017).

APPLICATIONS

I-motif structure has got several applications especially in the field of nanotechnology and analytical chemistry due to its pH driven “on/off” activity. This activity has been modulated for applications such as molecular switches, biosensors and nanomachines. Studies have been conducted upon its applications in pH analysis like in spectroscopy-based pH analysis, electrochemistry-based pH analysis and also in vivo pH analysis (Dong *et al*. 2014, Huang *et al*. 2014). Modi *et al.* (2009) developed the first of such pH sensor based on the FRET mechanism (named “I-switch”), which was able to sense and report pH changes along endosomal maturation both inside living cells in culture and in multicellular organism. pH sensors based on localized surface plasmon resonance (LSPR) and intracellular probe termed i-motif-based nanoflares have also been described (Wang *et al*. 2013, Huang *et al*. 2014). Gold nanoparticles (AuNPs) in conjugation with i-motif forming sequences have been efficiently developed as pH triggered drug delivery system (Song *et al*. 2013). Other applications in analytical

chemistry include glucose and pyruvic acid analysis, determination of diverse analytes, such as carbon nanotubes, Ag+ ions or proteins (Alba *et al*. 2013).

The interest in the study of this structure has increased again because of its possible future application as target for anticancer drug design and gene regulation processes. There is a low pH inside endosomes and owing to active metabolism of cancer cells there is acidic pH of the tumor microenvironment (Webb *et al*. 2011). These factors increase the interest for pH responsive systems for selective delivery and i-motif is an interesting pH sensitive DNA scaffold. Mesoporous silica nanoparticles with i-motif sequences as their cap have been used to generate a versatile delivery device having pores that open and close with changing the pH (Dong *et al*. 2014).

CONCLUSION AND FUTURE PERSPECTIVE

This review provides a brief introduction of i-motif structure, its biological relevance to cell biology, few of its applications and future possibilities. Recent confirmation of the existence of i-motif DNA structures *in vivo* have now opened new doors for further studies exploring its biological role and has also emphasized for its validation as a therapeutic target in cancer and other pathological conditions. Moreover, despite significant efforts being done towards the development of drug-like ligands of G-quadruplex structures, no suitable drug candidate has yet been developed (as a consequence of off-target binding, majority of known G-quadruplex-targeting ligands show remarkable toxicity). As compared to G-quadruplexes, i-motif DNA structures owing to their peculiar architecture can provide potentially higher degree of selective binding interactions, hence making them ostensible candidates for specific recognition processes. Still there are many challenges in understanding their structural stability and still a lot more further investigation is required on their interaction with ligands and proteins so that their full potential can be exploited for biomedical and other purposes. The recent studies have given a new insight to this non-canonical DNA form which will flourish in coming years.

ACKNOWLEDGEMENT

The authors are thankful to ICAR-Indian Veterinary Research Institute, Bareilly, Uttar Pradesh, India.

REFERENCES

Alba JJ, Sadurní A, Gargallo R (2016) Nucleic acid i-motif structures in analytical chemistry. Crit Rev Anal Chem 46(5): 443-54.

Amato J, Laccarino N, Randazzo A, Novellino E, Pagano B (2014) Noncanonical DNA secondary structures as drug targets: the prospect of the i-Motif. Chem Med Chem 9: 2026–2030.

Assi HA, Garavís M, Gonzalez C, Damha MJ (2018) i-Motif DNA: structural features and signi?cance to cell biology. Nucleic Acids Res 46(16): 8038–8056.

Berger I, Egli M, Rich A (1996) Inter-strand CH···O hydrogen bonds stabilizing four-stranded intercalated molecules: stereoelectronic effects of O4'in cytosine-rich DNA. Proc Natl Acad Sci 93(22): 12116-21.

Biffi G, Tannahill D, McCafferty J, Balasubramanian S (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. Nat Chem 5: 182–186.

Chen Y, Qu K, Zhao C, Wu L, Ren J, Wang J, Qu X (2012) Insights into the biomedical effects of carboxylated single-wall carbon nanotubes on telomerase and telomeres. Nat Commun 3: 1074.

Cui Y, Kong D, Ghimire C, Xu C, Mao H (2016) Mutually exclusive formation of G-quadruplex and i-motif is a general phenomenon governed by steric hindrance in duplex DNA. Biochemistry 55: 2291–2299.

Day HA, Pavlou P, Waller ZAE (2014) i-Motif DNA: Structure, stability and targeting with ligands. Bioorg Med Chem 22: 4407–4418.

Dexheimer TS, Carey SS, Zuohe S, Gokhale VM, Hu X, Murata LB, Maes EM, Weichsel A, Sun D, Meuillet EJ, Montfort WR, Hurley LH (2009) NM23-H2 may play an indirect role in transcriptional activation of c-myc gene expression but does not cleave the nuclease hypersensitive element III. Mol Cancer Ther 8: 1363-1377.

Dong Y, Yang Z, Liu D (2014) DNA Nanotechnology based on i-Motif structures. Acc Chem Res 47: 1853-1860.

Fleming AM, Ding Y, Rogers RA, Zhu J, Burton AD, Carlisle CB, Burrows CJ (2017) 4n-1 is a "Sweet Spot" in DNA i-Motif folding of 2'-Deoxycytidine homopolymers. J Am Chem Soc 139: 4682–4689.

Gehring K, Leroy JL, Gueron M (1993) A tetrameric DNA-structure with protonated cytosine-cytosine base-pairs. Nature 363: 561–565.

Henderson E, Hardin CC, Walk SK, Tinoco LJr., Blackburn EH (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. Cell 51: 899-908.

Henikoff S, Thakur J, Kasinathan S, Talbert PB (2017) Remarkable evolutionary plasticity of centromeric chromatin. Cold Spring Harb. Symp Quant Biol 82: 71–82.

Huang J, He Y, Yang X, Wang K, Ying L, Quan K, Yang Y, Yin B (2014) I-motif-based nano-flares for sensing pH changes in live cells. Chem Commun 50: 15768–15771.

Jafri MA, Ansari SA, Alqahtani MH, Shay JW (2016) Roles of telomeres and telomerase in cancer, and advances in telomerase targeted therapies. Genome Med 8: 69.

Kang HJ, Kendrick S, Hecht SM, Hurley LH (2014) The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. J Am Chem Soc 136: 4172–4185.

Kendrick S, Kang HJ, Alam MP, Madathil MM, Agrawal P, Gokhale V, Yang DZ, Hecht SM, Hurley LH (2014) The dynamic character of the BCL2 promoter i-Motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure. J Am Chem Soc 136: 4161–4171.

Mergny JL, Lacroix L, Han XG, Leroy JL, Helene C (1995) Intramolecular folding of pyrimidine oligodeoxynucleotides into an i-DNA Motif. J Am Chem Soc 117: 8887–8898.

Modi S, Swetha MG, Goswami D, Gupta GD, Mayor S, Krishnan Y (2009) A DNA nanomachine that maps spatial and temporal pH changes inside living cells. Nat Nanotechnol 4: 325–330.

Roy B, Talukder P, Kang HJ, Tsuen SS, Alam MP, Hurley LH, Hecht SM (2016) Interaction of individual structural domains of hnRNP LL with the BCL2 promoter i-motif DNA. J Am Chem Soc 138: 10950–10962.

Song L, Ho VH, Chen C, Yang Z, Liu D, Chen R, Zhou D (2013) Efficient, pH-triggered drug delivery using a pH-responsive DNA-conjugated gold nanoparticle. Adv Healthcare Mater 2(2): 275-280.

Sutherland C, Cui Y, Mao H, Hurley LH (2016) Mechanosensor mechanism controls the G-quadruplex/i-motif molecular switch in the MYC promoter NHE III1. J Am Chem Soc 138: 14138–14151.

Takahashi S, Brazier JA, Sugimoto N (2017) Topological impact of noncanonical DNA structures on Klenow fragment of DNA polymerase. Proc Natl Acad Sci USA 114: 9605–9610.

Wang C, Du Y, Wu Q, Xuan S, Zhou J, Song J, Shao F, Duan H (2013) Stimuli-responsive plasmonic core-satellite assemblies: I-motif DNA linker enabled intracellular pH sensing. Chem Commun 49: 5739-5741.

Webb BA, Chimenti M, Jacobson MP, Barber, DL (2011) Dysregulated pH: a perfect storm for cancer progression. Nat Rev Cancer 11: 671–677.

Wright EP, Huppert JL, Waller ZAE (2017) Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. Nucleic Acids Res 45: 2951–2959.

Yang B, Rodgers MT (2014) Base-pairing energies of proton-bound heterodimers of cytosine and modified cytosines: implications for the stability of DNA i-motif conformations. J Am Chem Soc 136(1): 282-90.

Zeraati M, Langley DB, Schofield P, Moye AL, Rouet R, Hughes WE, Bryan TM, Dinger ME, Christ D (2018) I-motif DNA structures are formed in the nuclei of human cells. Nat Chem10: 631–637.

Cite this article as: Nautiyal S, Rai V, Bhat S, Kumar R, Rather M M, Sankar M (2020) i-Motif DNA: Significance and future prospective. Explor Anim Med Res 10(1): 18-23.