nature chemistry

ARTICLES

https://doi.org/10.1038/s41557-018-0046-3

I-motif DNA structures are formed in the nuclei of human cells

Mahdi Zeraati[1,2], David B. Langley[1], Peter Schofield[1], Aaron L. Moye[3], Romain Rouet[1], William E. Hughes[1,2], Tracy M. Bryan[3], Marcel E. Dinger[1,2*] and Daniel Christ[1,2*]

Human genome function is underpinned by the primary storage of genetic information in canonical B-form DNA, with a second layer of DNA structure providing regulatory control. I-motif structures are thought to form in cytosine-rich regions of the genome and to have regulatory functions; however, in vivo evidence for the existence of such structures has so far remained elusive. Here we report the generation and characterization of an antibody fragment (iMab) that recognizes i-motif structures with high selectivity and affinity, enabling the detection of i-motifs in the nuclei of human cells. We demonstrate that the in vivo formation of such structures is cell-cycle and pH dependent. Furthermore, we provide evidence that i-motif structures are formed in regulatory regions of the human genome, including promoters and telomeric regions. Our results support the notion that i-motif structures provide key regulatory roles in the genome.

DNA has a well-known propensity to adopt alternative non-B-form conformations in vitro, including G-quadruplex (G4)[1] and intercalated motif (i-motif) structures[2] (Fig. 1a). Of the two structures, G4 DNA formed within guanine (G)-rich regions of the genome is by far the more studied. Computational analysis[3,4] and high-throughput G4 sequencing[5] have revealed that G4s are enriched in the regulatory regions of the genome. Furthermore, several studies have investigated regulatory functions of G4s in cellular pathways[6,7]. In spite of in silico and in vitro characterizations of G4s, the in vivo existence of them in human cells remained speculative until the visualization of these structures using an antibody fragment that recognizes G4 in a structure-specific manner[8].

In marked contrast, insights into the biological role of i-motif DNA are limited. This structure is formed via a stack of intercalating hemiprotonated C-neutral C base pairs (C^+:C), which are stabilized in slightly acidic pH (Fig. 1a)[9]. As C-rich regions are common within the human genome, i-motifs have been characterized in vitro, derived from telomeres[2,10], centromeres[11,12] and promoter regions of proto-oncogenes[13], using a range of biophysical techniques. These biophysical analyses demonstrated the formation of both intra- and intermolecular structures, with the overall stability of the structure dependent on the number of cytosines in the i-motif core, as well as the length and composition of loops[13–16].

Despite strong evidence that shows C-rich sequences can adopt i-motif structures in vitro, the in vivo existence of this four-stranded DNA structure in the human genome is still a matter of scientific debate. In particular, observations that in vitro formation of the i-motif structure is dependent on acidic conditions have raised questions concerning biological relevance[17,18]. On the other hand, the existence of i-motif DNA in vivo is supported by studies that demonstrated that this structure can form at physiological pH under conditions of molecular crowding[19–21] and negative superhelicity induced during transcription[22]. Recently, genomic sequences that form stable i-motif structures at neutral pH have been reported[23–25]. Furthermore, several studies have indicated that i-motif formation modulates replication[26] and transcription[27–31].

Taken together, while previous in vitro and indirect evidence has supported the notion that i-motif structures are formed in vivo, direct evidence has remained elusive. Here, we report the generation and characterization of a human antibody fragment that recognizes i-motif structures with high selectivity and affinity. We utilize this reagent to demonstrate the first direct visualization of i-motif structures in the nuclei of human cells.

Results

Selection and characterization of an antibody fragment specific for human telomere i-motif. We utilized the Garvan-2 human single-chain variable fragment (scFv) library[32–34] to select i-motif-specific antibody fragments. Three rounds of phage selections were performed to isolate binders to the human telomere i-motif (hTelo i-motif), a well-characterized i-motif[35] whose structure has been reported[2] (Fig. 1b and Supplementary Table 1). Binders were characterized using soluble fragment enzyme-linked immunosorbent assay (ELISA) followed by bio-layer interferometry (BLI) off-rate ranking, resulting in the identification of a lead candidate (which we termed iMab). A detailed protocol of the antibody selection procedure is provided in Supplementary Method 1 and Supplementary Table 2.

We next used BLI to determine the kinetics and affinity of the interaction between purified iMab and the hTelo i-motif. We observed a highly stable complex, characterized by a slow dissociation rate and a K_D in the low nanomolar range ($k_a = 3.2 \times 10^4\,s^{-1}\,M^{-1}$, $k_d = 1.9 \times 10^{-3}\,s^{-1}$, $K_D = 59\,nM$) (Fig. 1c). We next evaluated the specificity of iMab for the hTelo i-motif by ELISA. This analysis revealed highly specific binding, as indicated by the absence of binding to a wide range of control protein and nucleic acid antigens including double-stranded DNA, hairpin DNA, microRNA, streptavidin, neutravidin, hen egg-white lysozyme and neuropeptide Y (NPY) (Fig. 1d). In addition, no binding was observed to a mutant hTelo oligonucleotide (Fig. 1e), which lacks the i-motif structure, as indicated by circular dichroism (CD) spectroscopy (Fig. 1f). To investigate whether iMab is capable of binding to partially folded i-motif structures[36], we tested truncated constructs with either three or two

[1]Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research, Darlinghurst, Sydney, New South Wales, Australia. [2]St Vincent's Clinical School, Faculty of Medicine, University of New South Wales, Kensington, Sydney, New South Wales, Australia. [3]Children's Medical Research Institute, University of Sydney, Westmead, Sydney, New South Wales, Australia. *e-mail: m.dinger@garvan.org.au; d.christ@garvan.org.au

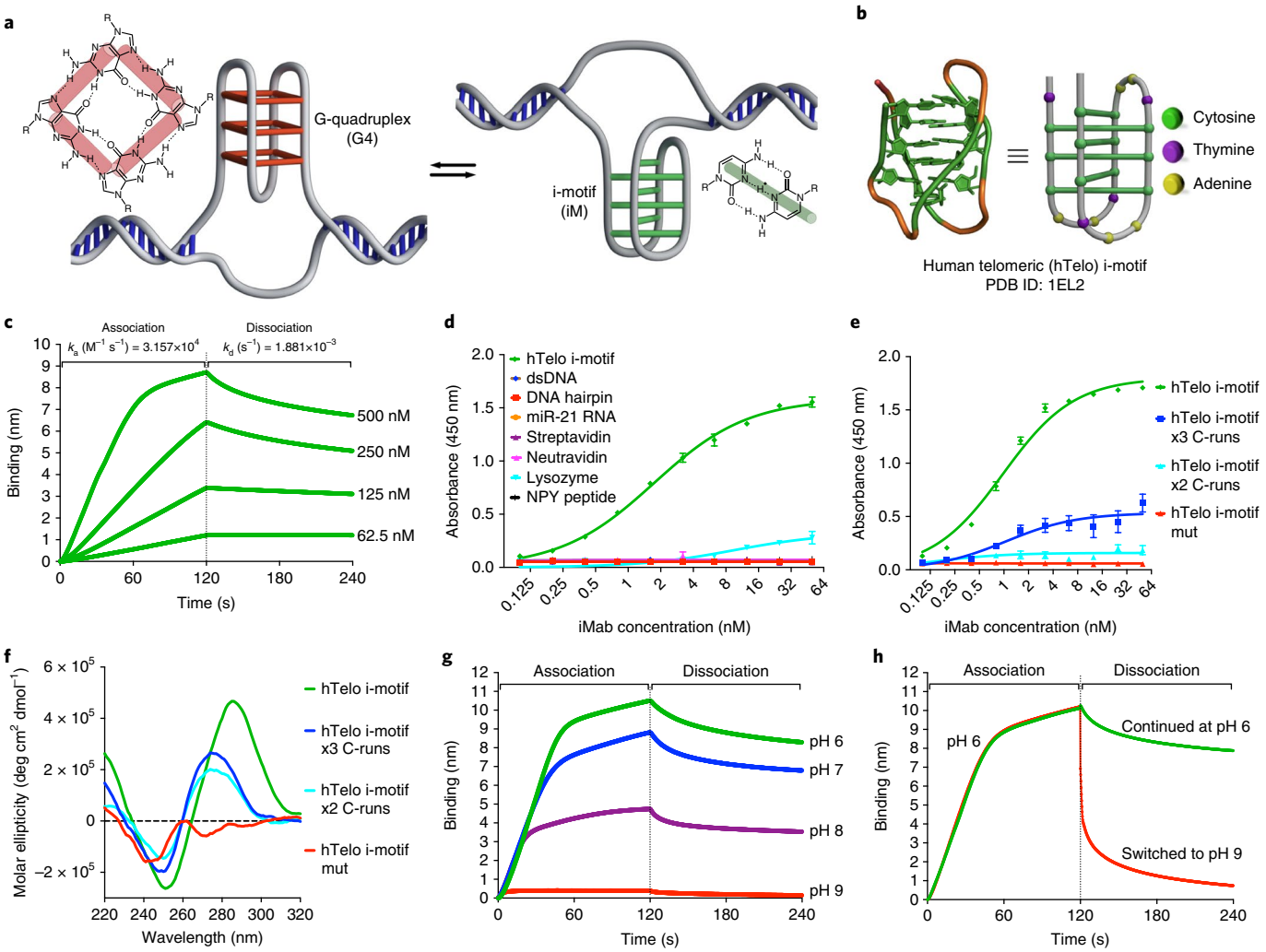

Fig. 1 | Characterization of an hTelo i-motif specific antibody fragment (iMab). a, Schematic of i-motif and G4 structures. **b**, Structure and schematic drawing of human telomeric (hTelo) i-motif (PDB: 1EL2). **c**, Bio-layer interferometry (BLI) analysis of binding of different concentrations of iMab to hTelo i-motif at pH 6 (association constant, $k_a = 3.2 \times 10^4\ s^{-1}\ M^{-1}$; dissociation constant, $k_d = 1.9 \times 10^{-3}\ s^{-1}$; equilibrium constant, $K_D = 59$ nM). **d**, Examination of iMab binding to control oligonucleotides and proteins (ELISA format). Plotted values are means ± s.d. of three replicates. **e**, Examination of iMab binding to full-length hTelo i-motif, as well as to truncated (x3 and x2 C-runs) and mutant species (ELISA format). Plotted values are means ± s.d. of three replicates. **f**, Circular dichroism (CD) spectra of full-length hTelo i-motif, truncated (x3 and x2 C-runs) and mutant species. **g**, BLI analysis of binding of iMab to hTelo i-motif at conditions at or above pH 6. **h**, BLI analysis of binding of iMab to hTelo i-motif at pH 6 and dissociation at pH 6 and 9, respectively.

tandem runs of cytosines. Indeed, these truncated hTelo i-motifs displayed reduced CD ellipticity compared to wild-type (wt) hTelo i-motif (Fig. 1f), suggesting an inability to form a fully folded i-motif structure. The two truncated constructs displayed either considerably reduced binding (three C-runs) or no detectable binding (two C-runs), indicating that recognition by iMab is highly structure dependent (Fig. 1e). In contrast, wt hTelo displayed a high degree of foldedness, as indicated by positive ellipticity at ~285 nm and negative ellipticity at ~255 nm (Fig. 1f). This was highly comparable to a previously reported CD spectrum for wt hTelo[35], and folded wt hTelo was recognized by iMab under all analysed conditions.

Next, the pH dependence of i-motif stability and iMab binding was investigated. In general, i-motif structures are considered most stable at acidic pH[37], reflecting the negative logarithmic acid dissociation constant (pK_a-) of cytosine and the requirement for the formation of hemiprotonated C^+:C pairs[35]. We thus hypothesized that pH and the binding of iMab should be inversely correlated, assuming that the antibody predominantly recognizes folded i-motif rather than unfolded species. We investigated this notion, using BLI and a constant concentration of iMab binding to the hTelo i-motif under conditions ranging from pH 6 to 9 (the antibody itself was highly stable under all analysed conditions; Supplementary Fig. 1). This experiment indeed indicated considerable reduction of antibody binding at higher pH and no detectable binding at pH 9 (Fig. 1g). In addition, antibody–i-motif complex formed at pH 6 could be readily dissociated when switching to pH 9 buffer conditions (Fig. 1h). Next we investigated the influence of iMab on the i-motif structure. CD spectroscopy was performed for the hTelo i-motif in the presence or absence of iMab across different pH values. This experiment revealed that the addition of iMab does not result in any wavelength shift irrespective of pH (Supplementary Fig. 2), implying that iMab binding does not detectably induce or alter the i-motif structure. Taken together, our results demonstrate that iMab recognizes the folded hTelo i-motif structure with high affinity and specificity.

iMab recognizes i-motifs in a structure-specific manner. In addition to telomere regions, C-rich sequences capable of adopting i-motif structures in vitro are commonly found in other regions

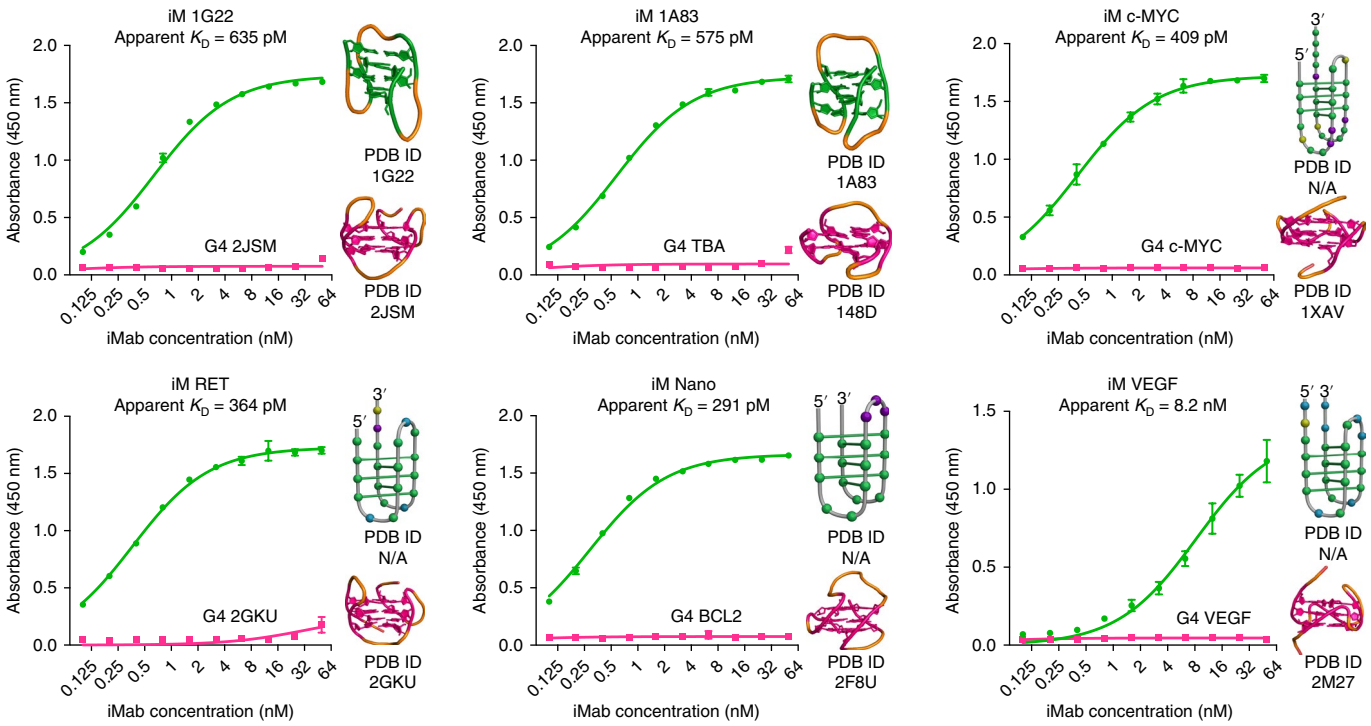

Fig. 2 | Evaluation of the specificity of iMab for different i-motif and G4 structures. iMab recognizes a wide range of i-motif DNA structures with high affinity; it does not detectably bind to G4 structures. Plotted values are means ± s.d. of three replicates. Structure and PDB identifiers are provided; predicted conformations are provided where structural information is unavailable. Green, cytosine; yellow, adenine; purple, thymine; blue, guanine.

throughout the genome. Several of these i-motifs have been biophysically characterized, including sequences from centromeres and the promoters of proto-oncogenes[13,18]. In addition to the hTelo i-motif structure described above, we evaluated the binding of iMab to six other well-defined intramolecular i-motif structures (Supplementary Table 1) by ELISA. These include 1G22[11] and 1A83[38], whose structures have been solved, and four other i-motifs, c-MYC[22], RET[39], VEGF[40] and Nano[41]. As depicted in Fig. 2, iMab has high apparent affinities for all of these i-motifs, ranging from approximately 290 picomolar to around 10 nanomolar. These results suggest iMab is an i-motif structure-specific antibody and capable of recognizing this structural motif with high affinity, irrespective of sequence diversity.

iMab differentiates i-motifs from G4s. Unlike i-motifs, for which only a small number of structures are available in the Protein Data Bank (PDB)[42], structures of many G4s have been reported; these structures demonstrate extensive structural and topological diversity[43]. Therefore, it is conceivable that i-motifs and G4s could display common epitopes. Moreover, in genomic DNA, i-motifs are often accompanied by complementary G-rich sequences, which can adopt G4 structures. For applications such as immunostaining, a high level of specificity and the absence of binding to G4s is therefore likely to be important. To evaluate iMab cross-reactivity we examined binding to six G4s of known structures, which were selected to represent a diverse range of conformations (Supplementary Table 1): promoter G4s c-MYC[44], BCL2[45], VEGF[46] and telomeric G4s 2JSM[47], 2GKU[48] and aptamer G4 TBA[49]. As outlined in Fig. 2, ELISA results show that iMab does not display any detectable binding to these molecules, suggesting that it is highly capable of differentiating between G4 and i-motif structures.

To confirm these results, binding was analysed by BLI using two additional antibodies as controls. As a negative control, we used the DP47DPK9 germline scFv antibody[32], from which iMab and the parental Garvan-2 scFv library were derived. While DP47DPK9 did not show any binding to i-motifs and G4s (Supplementary Fig. 3), iMab recognized i-motifs and differentiated them from G4s (Supplementary Fig. 3). As a second control, we utilized the widely used commercially available G4 quadruplex antibody BG4[8]. As expected, this antibody showed substantial binding to all analysed G4 structures (Supplementary Fig. 3); however, BG4 also displayed detectable cross-reactivity to i-motifs (Supplementary Fig. 3) and competed with iMab for binding to several i-motifs (Supplementary Fig. 4).

I-motif structures are formed in the nuclei of human cells. We then used iMab for indirect immunofluorescent staining of three different human cell lines. This staining revealed punctate foci in the nuclei of MCF7, U2OS and HeLa cell lines that were dependent on the presence of iMab (Fig. 3a). Next, we investigated whether DNase I digestion can affect the number of visualized foci. This experiment showed that DNase I treatment significantly decreased the number of foci in the nuclei (Fig. 3a,b; see Supplementary Method 2 for counting of the number of foci). Limited residual foci were observed after treatment, presumably due to nuclease-resistant DNA structures or shielding by proteins. Another possibility is the formation of RNA i-motifs that are not removed by DNase I digestion. A small decrease in the number of iMab foci after RNase A treatment supports this speculation (Fig. 3b). As a further control, we blocked iMab by pre-incubation with a fivefold molar excess of folded hTelo i-motif oligonucleotide before staining and observed a significant decline in the immunofluorescence signal (Fig. 3b). In addition, DP47DPK9 germline antibody staining did not show detectable punctate foci (Fig. 3b).

To validate that the detected foci correspond to i-motif structures, we transfected HeLa cells with folded i-motif and G4 oligonucleotides. Cells were transfected with 1, 2 and 4 µg of folded hTelo i-motif. Staining cells after 16 h using iMab showed an increase in the number of foci that directly correlated with the amount of transfected hTelo i-motif (Fig. 3c). The same experiment was performed

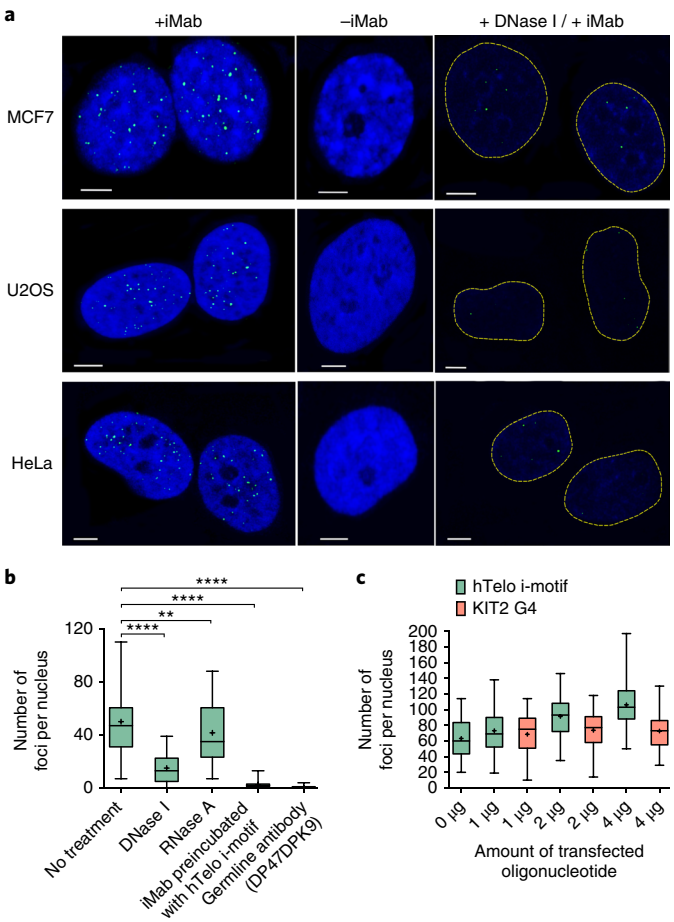

Fig. 3 | Visualization of i-motif structures in the nuclei of human cell lines. **a**, Imaging of iMab foci in the nuclei of MCF7, U2OS and HeLa cells using confocal microscopy. Nuclei were counterstained with DAPI. Yellow dashed lines in the DNase I treated samples indicate the borders of nuclei. Scale bars, 5 µm. **b**, Quantification of iMab foci across four different conditions and DP47DPK9 germline antibody foci in HeLa cells. DP47DPK9 and iMab foci of 200–300 nuclei were counted per condition. Boxes represent the 25th to 75th percentiles. Horizontal lines and '+' indicate medians and means, respectively. Whiskers indicate highest and lowest values of the results. $^{**}P < 0.01$, $^{****}P < 0.0001$; statistical significance was assessed by one-way ANOVA using Tukey's multiple comparisons test. **c**, Quantification of iMab foci after transfecting HeLa cells with either folded hTelo i-motif or folded KIT2 G4 oligonucleotides. Lipofectamine reagent was used for the transfection including control samples. iMab foci in 200–300 nuclei were counted per condition. Boxes represent the 25th to 75th percentiles. Horizontal lines and '+' indicate medians and means, respectively. Whiskers indicate highest and lowest values of the results.

for folded KIT2 G4[50] (Supplementary Table 1). We did not observe an increasing trend in the number of iMab foci on increasing the amount of transfected KIT2 G4 (Fig. 3c).

Cell-cycle-dependent formation of i-motif structures. To investigate whether i-motif formation is affected by cell-cycle progression, we used the iMab antibody to stain HeLa cells synchronized at G0/G1, G1/S boundary or early S phase. To analyse cell-cycle stages, we performed propidium iodide staining followed by flow cytometry. Synchronization at G0/G1 was reached by serum deprivation for 24 h. We observed the minimum number of foci at this phase. Incubation of the cells for 16 h in the presence of mimosine arrests them at the G1/S boundary[51]. Staining of these cells showed that i-motif formation is at the highest level at this phase of the cell cycle. Three hours after being released from the mimosine-induced blockade, when the cells synchronously begin S phase, the number of iMab foci was reduced compared to the G1/S boundary phase (Fig. 4a,b). Similar trends of i-motif formation during cell-cycle progression were observed in the MCF7 cell line (Supplementary Fig. 5).

pH-dependent formation of i-motif structures. We next examined how the formation of i-motif structures is affected by pH. As intracellular pH is inversely correlated with the concentration of CO_2 in the cell culture system[52,53], we hypothesized that we could investigate the effect of pH on i-motif formation by changing the CO_2 concentration. First, we incubated MCF7 cultures at 5% CO_2 until they reached ~70% confluency and then incubated the cultures in 2%, 5% or 8% CO_2 incubators for a further 2.5 h. When staining with iMab, a significant increase in focus number was observed with increasing CO_2 concentration (Fig. 5a,b). As expected, measurements of pH in the MCF7 culture medium indicated an inverse correlation with the CO_2 concentration (Fig. 5c). To verify that intracellular pH changes upon variation in CO_2 concentration, we used pHrodo Red as an intracellular pH indicator, a fluorogenic probe permeating the cell membrane (fluorescence intensity is inversely correlated with intracellular pH). As depicted in Fig. 5d, compared to 5% CO_2, we observed ~20% lower and ~30% higher fluorescence at 2% and 8% CO_2, respectively. These results indicate that intracellular pH can be decreased by incubating cells at higher CO_2 concentrations, resulting in increased i-motif structure formation. To confirm this result, we repeated the experiments using HEK 293T cells and pHrodo Green, a different intracellular pH indicator, to exclude cell line and reagent-specific effects. Similar trends of pH- and CO_2-dependent i-motif formation were observed (Supplementary Fig. 6).

Evidence for i-motif formation in regulatory regions. G/C-rich repeats capable of forming in vitro G4 and i-motif structures, respectively, have been identified in human telomeres. We explored the in vivo formation of i-motif structures in the telomere region by co-staining U2OS cells using iMab and a TRF2 (telomere repeat-binding factor 2) antibody. TRF2 is a negative regulator of telomere length and localizes predominantly at telomeres[54]. Using confocal microscopy, we observed some co-localization of signals in U2OS cells stained with iMab and TRF2 antibodies (Fig. 6a). The structures recognized by these reagents are probably beyond the resolving power of conventional microscopy, even using deconvolution algorithms. Therefore, to confirm the likelihood of the co-localization, we used stimulated emission depletion (STED) microscopy, which can provide super-resolution (<50 nm in x–y) three-dimensional images. Using STED microscopy, we were able to detect the co-localization of iMab and TRF2 foci (Supplementary Fig. 7). Our data suggest that i-motif structures are likely to exist at TRF2 foci on telomeres; however, a considerable number also form outside telomere regions.

It has been suggested that C-rich sequences capable of forming stable i-motif structures at physiological pH are enriched in regulatory regions of the human genome[23,24]. Among these sequences, i-motifs in the promoter regions of proto-oncogenes[13] (such as *BCL2*[27,28,55,56]) are arguably the most well-studied. We investigated the in vivo formation of i-motifs in promoter regions by co-staining U2OS cells with iMab and an anti-E12/E47 antibody. The E12/E47 transcription factors are encoded by the *E2A* gene and bind to the enhancer (E) box[57]. We observed several concurrences between iMab and E12/E47 foci (Fig. 6b). Our result further supports the notion that i-motif structures form in promoter regions and modulate transcription.

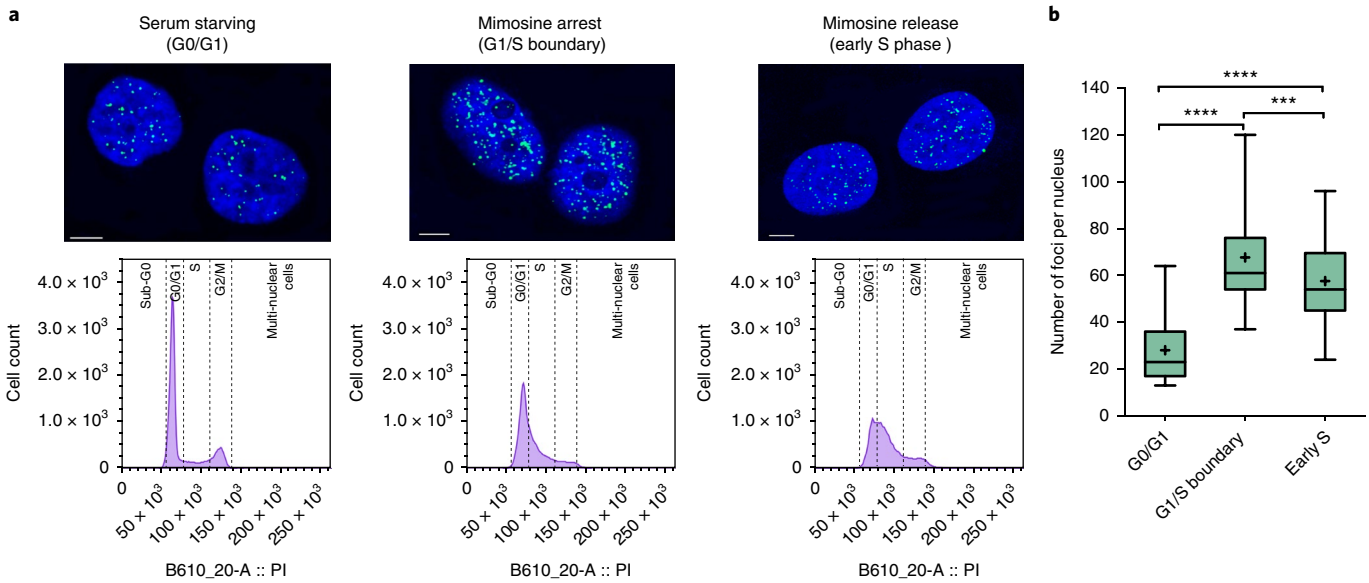

Fig. 4 | Visualization and quantification of i-motif structures during cell-cycle progression in HeLa cells. a, Imaging of iMab foci in the nuclei of HeLa cells using confocal microscopy. Cells were synchronized at the G0/G1, G1/S boundary and early S phases. Nuclei were counterstained with DAPI. Scale bars, 5 µm. Flow cytometry graphs outline the population of synchronized and propidium iodide stained cells in each phase. **b**, Quantification of the iMab foci in the nuclei of HeLa cells synchronized at the G0/G1, G1/S boundary and early S phases. iMab foci of 200–300 nuclei were counted per condition. Boxes represent the 25th to 75th percentiles. Horizontal lines and '+' indicate medians and means, respectively. Whiskers indicate highest and lowest values of the results. ***$P < 0.001$, ****$P < 0.0001$. Statistical significance was assessed by one-way ANOVA using Tukey's multiple comparisons test.

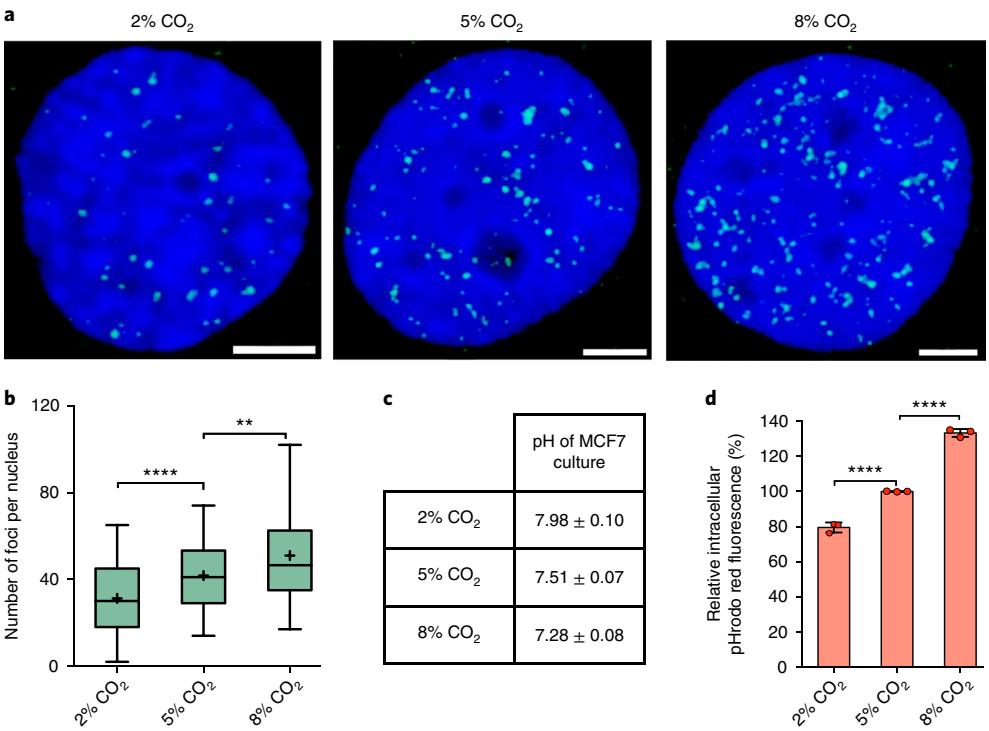

Fig. 5 | Investigation of the effect of pH variation on the formation of i-motif structures in MCF7 cells. a, Imaging of iMab foci in the nuclei of MCF7 cells after incubation at different CO_2 concentrations for 2.5 h using confocal microscopy. Confocal images show iMab foci in the nuclei of MCF7 cells. Nuclei were counterstained with DAPI. Scale bars, 3 µm. **b**, Quantification of iMab foci in the nuclei of MCF7 cells after 2.5 h variation of CO_2 concentration. iMab foci of 200–300 nuclei were counted per condition. Boxes represent 25th to 75th percentiles. Horizontal lines and '+' indicate medians and means, respectively. Whiskers indicate highest and lowest values of the results. **$P < 0.01$, ****$P < 0.0001$; statistical significance was assessed by one-way ANOVA using Tukey's multiple comparisons test. **c**, pH values of MCF7 cultures after 2.5 h variation in CO_2 concentration. **d**, Relative intracellular pHrodo Red fluorescence across different CO_2 concentrations in MCF7 cells. Error bars represent mean ± s.e.m. of three replicates. ****$P < 0.0001$; statistical significance was assessed by one-way ANOVA using Tukey's multiple comparisons test.

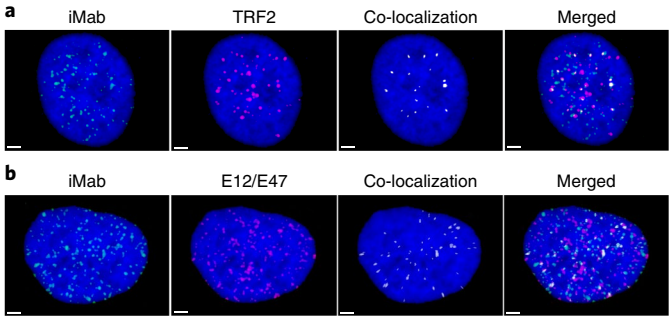

Fig. 6 | Evidence of i-motifs formation in regulatory regions. a, Imaging of iMab and TRF2 co-localization at telomeres in the nucleus of an U2OS cell using confocal microscopy. **b**, Imaging of iMab and E12/E47 co-localization at promoters in the nucleus of an U2OS cell using confocal microscopy. Nuclei were counterstained with DAPI. Scale bars, 3 µm. Imaris software was used to analyse co-localizations.

Discussion

By generating an i-motif-specific antibody fragment and utilizing it for immunofluorescent staining, we provide the first direct evidence for the presence of i-motif structures in the nuclei of human cells. We demonstrate dynamic i-motif formation in vivo, dependent on cell-cycle progression and pH variation. In addition, we describe evidence for the formation of i-motif structures within regulatory regions of the human genome.

By providing a highly specific probe, the research outlined here overcomes a longstanding technical limitation in the i-motif field. Much as the development of anti-G4 antibody fragments has revolutionized this field over the last 15 years[8,58], the availability of high-quality i-motif-specific reagents, which differentiate these structures from other secondary structures such as G4s, has the potential to provide transformative advances in our understanding of i-motif functions in the cell. Our observation that commonly used anti-G4 antibodies cross-react with i-motif DNA also raises implications for the G4 quadruplex field.

One of the main sources of uncertainty regarding the biological relevance of i-motif structures is the fact that such structures had so far been predominately observed at acidic pH below the physiological range[9]. Given the high sensitivity of iMab detection (Supplementary Fig. 8), we were able to detect i-motifs within a pH range of 6–8 (Fig. 1g). Our results thus indicate that i-motif DNA structures can indeed form under physiological conditions, presumably further supported by cellular features such as negative superhelicity[22], epigenetic modifications[59] and molecular crowding[19–21].

Our results also suggest that the formation of i-motif structures is highly cell-cycle specific. More specifically, using iMab staining, we observed that the highest level of i-motif formation occurs in late G1 phase, which is characterized by high levels of transcription and cellular growth (Fig. 4a,b and Supplementary Fig. 5). This observation agrees with the findings described for the regulatory role of i-motif structures at the promoters of several proto-oncogenes[27–29,55,56], further supporting the notion that i-motifs may act as scaffolds for the binding of transcription factors during transcription. This behaviour is markedly different to G4 formation, which occurs predominately during the S phase[8]. In contrast, we observed a reduction in the number of iMab stained foci in early S phase, indicating that i-motif structures are resolved during replication. The difference between i-motif and G4 formation during cell-cycle progression agrees with recent findings demonstrating that these structures are often mutually exclusive[60] and play opposite roles in the regulation of gene expression[29].

Compared to G4s, which are significantly more stable under physiological conditions, the transient and pH-dependent formation of i-motif structures has considerably slowed research in the i-motif field. However, it is possible that the transient nature of the i-motif structure, far from being a biophysical artefact, may actually mediate important regulatory roles; more specifically, variation in local conditions such as pH may be modulating function by controlling i-motif stability, for instance within promoter and telomere regions.

We conclude that our results support the notion that C-rich sequences within the genome form i-motif structures in the nuclei of human cells and control regulatory functions. Therefore, it is conceivable that any changes, such as mutations that alter i-motif stability and conformation, can affect their regulatory role, as has been previously reported for other DNA and RNA structures[61,62]. Our results thereby provide a core foundation for future studies exploring the biological role of this common genomic DNA structure, and for validation as a therapeutic target in cancer and other pathological conditions.

Methods

Material and methods are described in the Supplementary Information.

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability. Data sets generated during the study and antibody preparations are available from the corresponding authors upon reasonable request.

Received: 29 June 2017; Accepted: 14 March 2018;
Published online: 23 April 2018

References

1. Parkinson, G. N., Lee, M. P. & Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **417**, 876–880 (2002).
2. Phan, A. T., Gueron, M. & Leroy, J. L. The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.* **299**, 123–144 (2000).
3. Huppert, J. L. & Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **33**, 2908–2916 (2005).
4. Maizels, N. & Gray, L. T. The G4 genome. *PLoS Genet.* **9**, e1003468 (2013).
5. Chambers, V. S. et al. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* **33**, 877–881 (2015).
6. Bochman, M. L., Paeschke, K. & Zakian, V. A. DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.* **13**, 770–780 (2012).
7. Rhodes, D. & Lipps, H. J. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* **43**, 8627–8637 (2015).
8. Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182–186 (2013).
9. Day, H. A., Pavlou, P. & Waller, Z. A. i-Motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.* **22**, 4407–4418 (2014).
10. Leroy, J. L., Gueron, M., Mergny, J. L. & Helene, C. Intramolecular folding of a fragment of the cytosine-rich strand of telomeric DNA into an i-motif. *Nucleic Acids Res.* **22**, 1600–1606 (1994).
11. Nonin-Lecomte, S. & Leroy, J. L. Structure of a C-rich strand fragment of the human centromeric satellite III: a pH-dependent intercalation topology. *J. Mol. Biol.* **309**, 491–506 (2001).
12. Garavis, M., Escaja, N., Gabelica, V., Villasante, A. & Gonzalez, C. Centromeric alpha-satellite DNA adopts dimeric i-motif structures capped by AT Hoogsteen base pairs. *Chemistry* **21**, 9816–9824 (2015).
13. Brooks, T. A., Kendrick, S. & Hurley, L. Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.* **277**, 3459–3469 (2010).
14. Gurung, S. P., Schwarz, C., Hall, J. P., Cardin, C. J. & Brazier, J. A. The importance of loop length on the stability of i-motif structures. *Chem. Commun.* **51**, 5630–5632 (2015).
15. Benabou, S. et al. Understanding the effect of the nature of the nucleobase in the loops on the stability of the i-motif structure. *Phys. Chem. Chem. Phys.* **18**, 7997–8004 (2016).
16. Fujii, T. & Sugimoto, N. Loop nucleotides impact the stability of intrastrand i-motif structures at neutral pH. *Phys. Chem. Chem. Phys.* **17**, 16719–16722 (2015).
17. Jin, K. S. et al. pH-dependent structures of an i-motif DNA in solution. *J. Phys. Chem. B* **113**, 1852–1856 (2009).
18. Benabou, S., Avino, A., Eritja, R., Gonzalez, C. & Gargallo, R. Fundamental aspects of the nucleic acid i-motif structures. *RSC Adv.* **4**, 26956–26980 (2014).

19. Cui, J., Waltman, P., Le, V. H. & Lewis, E. A. The effect of molecular crowding on the stability of human c-MYC promoter sequence I-motif at neutral pH. *Molecules* **18**, 12751–12767 (2013).
20. Rajendran, A., Nakano, S. & Sugimoto, N. Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun.* **46**, 1299–1301 (2010).
21. Li, H., Hai, J., Zhou, J. & Yuan, G. The formation and characteristics of the i-motif structure within the promoter of the c-myb proto-oncogene. *J. Photochem. Photobiol.* **B 162**, 625–632 (2016).
22. Sun, D. & Hurley, L. H. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *J. Med Chem.* **52**, 2863–2874 (2009).
23. Wright, E. P., Huppert, J. L. & Waller, Z. A. E. Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.* **45**, 2951–2959 (2017).
24. Fleming, A. M. et al. 4n–1 is a 'sweet spot' in DNA i-motif folding of 2′ -deoxycytidine homopolymers. *J. Am. Chem. Soc.* **139**, 4682–4689 (2017).
25. Mir, B. et al. Prevalent sequences in the human genome can form mini i-motif structures at physiological pH. *J. Am. Chem. Soc.* **139**, 13985–13988 (2017).
26. Takahashi, S., Brazier, J. A. & Sugimoto, N. Topological impact of noncanonical DNA structures on Klenow fragment of DNA polymerase. *Proc. Natl Acad. Sci. USA* **114**, 9605–9610 (2017).
27. Kang, H. J., Kendrick, S., Hecht, S. M. & Hurley, L. H. 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 (2014).
28. Kendrick, S. et al. 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 (2014).
29. Sutherland , C., Cui, Y., Mao, H. & Hurley , L. H. Mechanosensor mechanism controls the G-quadruplex/i-motif molecular switch in the *MYC*promoter NHE III1. *J. Am. Chem. Soc.* **138**, 14138–14151 (2016).
30. Kaiser, C. E. et al. Insight into the complexity of the i-motif and G-quadruplex DNA structures formed in the KRAS promoter and subsequent drug-induced gene repression. *J. Am. Chem. Soc.* **139**, 8522–8536 (2017).
31. Brown, R. V. et al. The consequences of overlapping G-quadruplexes and i-motifs in the platelet-derived growth factor receptor beta core promoter nuclease hypersensitive element can explain the unexpected effects of mutations and provide opportunities for selective targeting of both structures by small molecules to downregulate gene expression. *J. Am. Chem. Soc.* **139**, 7456–7475 (2017).
32. Dudgeon, K. et al. General strategy for the generation of human antibody variable domains with increased aggregation resistance. *Proc. Natl Acad. Sci. USA* **109**, 10879–10884 (2012).
33. Rouet, R., Dudgeon, K. & Christ, D. Generation of human single domain antibody repertoires by Kunkel mutagenesis. *Methods Mol. Biol.* **907**, 195–209 (2012).
34. Rouet, R., Dudgeon, K., Christie, M., Langley, D. & Christ, D. Fully human V_H single domains that rival the stability and cleft recognition of camelid antibodies. *J. Biol. Chem.* **290**, 11905–11917 (2015).
35. Manzini, G., Yathindra, N. & Xodo, L. E. Evidence for intramolecularly folded i-DNA structures in biologically relevant CCC-repeat sequences. *Nucleic Acids Res.* **22**, 4634–4640 (1994).
36. Dhakal, S., Lafontaine, J. L., Yu, Z., Koirala, D. & Mao, H. Intramolecular folding in human ILPR fragment with three C-rich repeats. *PLoS One* **7**, e39271 (2012).
37. Kim, B. G. & Chalikian, T. V. Thermodynamic linkage analysis of pH-induced folding and unfolding transitions of i-motifs. *Biophys. Chem.* **216**, 19–22 (2016).
38. Han, X., Leroy, J. L. & Gueron, M. An intramolecular i-motif: the solution structure and base-pair opening kinetics of d(5mCCT3CCT3ACCT3CC). *J. Mol. Biol.* **278**, 949–965 (1998).
39. Guo, K. et al. Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene. *J. Am. Chem. Soc.* **129**, 10220–10228 (2007).
40. Guo, K., Gokhale, V., Hurley, L. H. & Sun, D. Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. *Nucleic Acids Res.* **36**, 4598–4608 (2008).
41. Li, T. & Famulok, M. I-motif-programmed functionalization of DNA nanocircles. *J. Am. Chem. Soc.* **135**, 1593–1599 (2013).
42. Berman, H. M. et al. The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
43. Karsisiotis, A. I., O'Kane, C. & Webba da Silva, M. DNA quadruplex folding formalism—a tutorial on quadruplex topologies. *Methods* **64**, 28–35 (2013).
44. Ambrus, A., Chen, D., Dai, J., Jones, R. A. & Yang, D. Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry* **44**, 2048–2058 (2005).
45. Dai, J., Chen, D., Jones, R. A., Hurley, L. H. & Yang, D. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Res.* **34**, 5133–5144 (2006).
46. Agrawal, P., Hatzakis, E., Guo, K., Carver, M. & Yang, D. Solution structure of the major G-quadruplex formed in the human VEGF promoter in K^+: insights into loop interactions of the parallel G-quadruplexes. *Nucleic Acids Res.* **41**, 10584–10592 (2013).
47. Phan, A. T., Kuryavyi, V., Luu, K. N. & Patel, D. J. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K^+ solution. *Nucleic Acids Res.* **35**, 6517–6525 (2007).
48. Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L. & Patel, D. J. Structure of the human telomere in K^+ solution: an intramolecular (3+1) G-quadruplex scaffold. *J. Am. Chem. Soc.* **128**, 9963–9970 (2006).
49. Schultze, P., Macaya, R. F. & Feigon, J. Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.* **235**, 1532–1547 (1994).
50. Kuryavyi, V., Phan, A. T. & Patel, D. J. Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter. *Nucleic Acids Res.* **38**, 6757–6773 (2010).
51. Jackman, J. & O'Connor, P. in *Current Protocols in Cell Biology* Ch. 8 (John Wiley and Sons, 1998).
52. Simon, S., Roy, D. & Schindler, M. Intracellular pH and the control of multidrug resistance. *Proc. Natl Acad. Sci. USA* **91**, 1128–1132 (1994).
53. Huang, Z. & Huang, Y. The change of intracellular pH is involved in the cisplatin-resistance of human lung adenocarcinoma A549/DDP cells. *Cancer Invest.* **23**, 26–32 (2005).
54. Smogorzewska, A. et al. Control of human telomere length by TRF1 and TRF2. *Mol. Cell Biol.* **20**, 1659–1668 (2000).
55. Kendrick, S., Akiyama, Y., Hecht, S. M. & Hurley, L. H. The i-motif in the bcl-2 P1 promoter forms an unexpectedly stable structure with a unique 8:5:7 loop folding pattern. *J. Am. Chem. Soc.* **131**, 17667–17676 (2009).
56. Roy, B. et al. Interaction of individual structural domains of hnRNP LL with the BCL2 promoter i-motif DNA. *J. Am. Chem. Soc.* **138**, 10950–10962 (2016).
57. Aronheim, A., Shiran, R., Rosen, A. & Walker, M. D. The E2A gene product contains two separable and functionally distinct transcription activation domains. *Proc. Natl Acad. Sci. USA* **90**, 8063–8067 (1993).
58. Schaffitzel, C. et al. In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylonychia lemnae* macronuclei. *Proc. Natl Acad. Sci. USA* **98**, 8572–8577 (2001).
59. Xu, B., Devi, G. & Shao, F. Regulation of telomeric i-motif stability by 5-methylcytosine and 5-hydroxymethylcytosine modification. *Org. Biomol. Chem.* **13**, 5646–5651 (2015).
60. Cui, Y., Kong, D., Ghimire, C., Xu, C. & Mao, H. Mutually exclusive formation of G-quadruplex and i-motif is a general phenomenon governed by steric hindrance in duplex DNA. *Biochemistry* **55**, 2291–2299 (2016).
61. Wells, R. D. Non-B DNA conformations, mutagenesis and disease. *Trends Biochem. Sci.* **32**, 271–278 (2007).
62. Zeraati, M. et al. Cancer-associated noncoding mutations affect RNA G-quadruplex-mediated regulation of gene expression. *Sci. Rep.* **7**, 708 (2017).

Acknowledgements

The authors thank S. Thompson (Leica Microsystems Australia) and P. Young (University of Sydney) for access to the Leica 3XSTED instrument. This work was supported by Program Grants 1113904, Project Grant 1148051, Development Grants 1113790 and 1076356 and Fellowship 105146 from the National Health and Medical Research Council (NHMRC) and Discovery Grants 160104915 and 140103465 from the Australian Research Council (ARC).

Authors contributions

M.Z., M.E.D. and D.C. conceived the project. M.Z. designed and performed the experiments. D.B.L. and R.R. contributed phage display experiments. P.S. contributed in vitro characterization of the antibody. A.L.M. and T.M.B. contributed biophysical studies. W.E.H. contributed microscopy. M.Z., M.E.D. and D.C. analysed the data and wrote the paper. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41557-018-0046-3.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to M.E.D. or D.C.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Prof. Daniel Christ & Prof. Marcel Dinger

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Experimental design

1. Sample size

Describe how sample size was determined.

iMab foci of 300 - 400 nuclei were counted per condition for each biological replicate. Means shown for clarity purposes.

2. Data exclusions

Describe any data exclusions.

No data were excluded from the analyses.

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

For each experiment and condition, at least three independent technical replicates were performed for each of the two biological replicates. All observations reported in the manuscript were reproducible.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

N/A

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

N/A

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ Test values indicating whether an effect is present *Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes noted.*
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	1) Number of foci were counted using FIJI as described in the supplementary information. 2) Imaris version 8 was used to analyze the colocalization assay. 3) Statistical analyses was performed using GraphPad Prism version 7. 4) Global fitting curves in the BLI experiment were analyzed using BLItz Pro version 1.2.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.	iMab antibody fragment generated at the Garvan Institute would be available upon request for research applications.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	1) Anti-myc horseradish peroxidase (HRP) antibody (CMYC-45P, ICL) 2) Anti-FLAG antibody (14793S, Cell Signaling Technology) 3) Anti-rabbit Alexa 488-cojugated antibody (711-546-152, Jackson ImmunoResearch) 4) Anti-TRF2 antibody (ab13579, Abcam) 5) Anti-mouse Alexa 488-cojugated antibody (715-545-150, Jackson ImmunoResearch) 6) Anti-rabbit Alexa 594-cojugated antibody (A-11037, ThermoFisher) 7) BG4 antibody (Ab00174-1.1, Absolute Antibody) 8) Anti-E12/E47 antibody (sc-133075, Santa Cruz Biotechnology) 9) Anti-mouse Alexa 594-conjugated antibody (A11005, ThermoFisher)

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	MCF7: Michigan Cancer Foundation HeLa: Garvan Institute of Medical Research HEK 293T: Garvan Institute of Medical Research U2OS: Garvan Institute of Medical Research
b. Describe the method of cell line authentication used.	MCF7: validated by Cellbank Australia using short tandem repeat (STR) profiling HeLa: not done HEK 293T: validated by Cellbank Australia using STR profiling U2OS: not done
c. Report whether the cell lines were tested for mycoplasma contamination.	All cell lines were tested and are negative for mycoplasma
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	Cell lines used in this study were chosen as models and results did not show any cell line dependency.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.	N/A

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	N/A