

<https://doi.org/10.1038/s42004-024-01387-w>

# Modulation of Nrf2 expression by targeting i-motif DNA

E. F. Warner<sup>1</sup>, D. Guneri<sup>2</sup>, M. A. O'Connell<sup>1</sup>, C. J. MacDonald<sup>1</sup> & Z. A. E. Waller<sup>2</sup>✉

Nuclear factor (erythroid-derived 2)-like 2 (Nrf2) is a key regulator of cell detoxification, which maintains homeostasis in healthy cells and promotes chemoresistance in cancer cells. Controlling the expression of this transcription factor is therefore of great interest. There are many compounds that have been shown to induce *Nrf2* expression, but ligands that can inhibit Nrf2 are scant. Herein we characterise an i-motif-forming sequence downstream of the *Nrf2* promoter, which we hypothesised may regulate the expression of the gene. The Nrf2 i-motif was found to be stable at near-physiological conditions. We identified small molecule ligands that interact with this i-motif structure and one significantly upregulated Nrf2 mRNA expression, and one ligand reduced Nrf2 mRNA expression in human cancer cells. This is the first example of controlling the promoter of *Nrf2* by targeting DNA structures and offers an alternative mode of action for the development of compounds to improve the chemotherapeutic responsiveness of existing treatments for cancer.

Nuclear factor (erythroid-derived 2)-like 2 (Nrf2) is a transcription factor that regulates the expression of proteins serving a wide range of intracellular functions, such as detoxification and reduction of reactive oxygen species, which have anti-inflammatory and cytoprotective properties<sup>1</sup>. Global mapping of binding sites for Nrf2 has validated over 240 genes as targets for Nrf2 in cell survival response<sup>2</sup>. Nrf2 has been shown to have potential as a chemopreventive by effects on detoxification, response to oxidative stress and inhibiting DNA damage and inflammation<sup>3</sup>. It is also involved in tumour promotion through not only chemoresistance but in many of the hallmarks of cancer and potential drug resistance in tumours<sup>4</sup>. The prevalence of Nrf2 in the hallmarks of cancer strongly suggests that targeting this transcription factor could be a good therapeutic approach<sup>5</sup>.

Under normal conditions, Nrf2 is ubiquitinated and degraded under the regulation of Kelch-like ECH-associated protein 1 (Keap1). Upon activation, such as by an inflammatory or oxidative stimulus, Nrf2 accumulates in the nucleus and subsequently activates transcription through binding to the antioxidant response element (ARE) in the *Nrf2* promoter region alongside small Maf proteins<sup>1</sup>, regulating the expression of cytoprotective enzymes, such as haem oxygenase-1 (HO-1) and NAD(P)H dehydrogenase quinone 1 (NQO1). Many ligands have been shown to disrupt the interaction of Nrf2/Keap1<sup>6</sup>, to the best of our knowledge none have reported the therapeutic potential of controlling the promoter of *Nrf2* directly. It has been shown that ligands that target DNA sequences capable of forming secondary structures (e.g. G-quadruplex and i-motif) present in the promoter regions of several genes, can modulate gene expression<sup>7–9</sup>.

Given its dualistic nature and pathophysiological importance, Nrf2 is therefore an interesting target for this mechanism.

We have previously identified a G-quadruplex forming sequence on the leading strand within the *Nrf2* promoter sequence<sup>10</sup>. Others have shown that a G-quadruplex in the Nrf2 mRNA 5' untranslated region regulates de novo Nrf2 protein translation under oxidative stress<sup>11,12</sup>. Despite this, the putative i-motif structure in the promoter region of the gene, which would form opposite the DNA G-quadruplex, has not been investigated at all. i-Motifs are non-canonical quadruplex secondary DNA structures defined as two parallel stranded duplexes held in an antiparallel manner by the intercalation of hemi-protonated cytosine–cytosine<sup>+</sup> base pairs<sup>13</sup>. i-Motifs require hemi-protonation of cytosines to stabilise in vitro, however it has been demonstrated that this can occur at neutral pH<sup>14</sup> in conditions such as molecular crowding<sup>15</sup>. The presence of i-motif structures in cells has been shown by use of fluorescent antibodies and in-cell NMR experiments<sup>16–18</sup>, as well as whole genome methods<sup>19</sup>. This has increased the importance of understanding the functional role of these DNA structures and how these can be modulated for therapeutic purposes, particularly for targeting proteins where intervention cannot be achieved by other means<sup>20</sup>.

Here we characterise the i-motif forming sequence within the promoter region of the *Nrf2* gene by use of circular dichroism (CD) spectroscopy, UV spectroscopy, and <sup>1</sup>H NMR. Further, we present ligands identified through a high-throughput fluorescent indicator displacement (FID) assay and determine the effect of these ligands on carcinoma cell viability and their bioactivity on *Nrf2* gene expression.

<sup>1</sup>School of Chemistry, Pharmacy and Pharmacology, University of East Anglia, Norwich, Norfolk, UK. <sup>2</sup>UCL School of Pharmacy, London, UK.

✉ e-mail: [z.waller@ucl.ac.uk](mailto:z.waller@ucl.ac.uk)

## Results and Discussion

### *Nrf2* promoter sequence forms i-motif at near physiological pH conditions

The *Nrf2* promoter sequence has been shown to contain a G-quadruplex at –606 base pairs from the transcription start site and was identified as a potential quadruplex forming sequence<sup>10</sup>. The sequence, which we will refer to as Nrf2G, denoting it as the G-rich sequence 5'-GGG-AAG-GGA-GCA-AGG-GCG-GGA-GGG-3', is close to putative transcription factor binding sites in the *Nrf2* promoter<sup>21</sup>. To examine the i-motif we used the reverse complement, termed Nrf2C: 5'-CCC-TCC-CGC-CCT-TGC-TCC-CTT-CC-3'.

Circular dichroism (CD) was utilised to determine the transitional pH ( $pH_T$ ) of the oligonucleotide sequence upstream of the *Nrf2* transcription start site (Fig. 1A). Samples of Nrf2C (10  $\mu$ M DNA) were annealed in buffer containing 10 mM sodium cacodylate, 100 mM potassium chloride, at pH 4.0–8.0. i-Motif structure was determined by previously characterised parameters: a positive peak at ~288 nm and a negative peak at ~255 nm<sup>22</sup>. The  $pH_T$  of this i-motif forming Nrf2 sequence was determined to be pH 6.7 (Fig. 1A). The sequence is a relatively short i-motif forming sequence [5'-CCC-TCC-CGC-CCT-TGC-TCC-CTT-CCC-3'] with a tract length of three cytosines. Compared to other i-motif forming sequences, this is relatively stable for a tract length of three cytosines. For example, the human telomeric sequence has a  $pH_T$  of 6.3<sup>23</sup>, and the i-motif from the promoter region of the hras gene a  $pH_T$  of 5.9<sup>24</sup>.

UV spectroscopy was utilised to determine the thermal stability ( $T_m$ ) at pH 5.0, 5.5, 6.0, 6.5, 7.0 and 7.5 (Supplementary Fig. S1 and Supplementary Table S1). The  $T_m$  of the structure at pH 6.5 was  $27 \pm 0.4$  °C, making these conditions suitable for further investigation at room temperature. There were no significant differences between melting and annealing experiments, indicating fast folding and unfolding kinetics. This is in contrast to many i-motif forming sequences examined previously under the same conditions, where hysteresis was observed<sup>14,23</sup>. UV spectroscopy was further utilised to determine the thermal difference spectra (TDS) of Nrf2 (Fig. 1B). This measurement was used to characterise the dominant folded species of the Nrf2 secondary structures across a pH range. At pH 5.0–6.5, the TDS of Nrf2 demonstrated a positive peak at 240 nm and a negative peak at 295 nm, characteristic of folded i-motif structure<sup>25</sup>. Even at pH 7.0 the spectrum was consistent with i-motif, but the appearance was significantly different to the other pHs examined.

<sup>1</sup>H NMR was used to verify the formation of this i-motif structure at pH 6.5 (Supplementary Fig. S2), where imino proton signals can be observed at 15.5 ppm at 298 K, characteristic of the C<sup>+</sup>–C base pairs in i-motif<sup>26</sup>. Further imino signals present at 12.5–13 ppm indicate additional Watson-Crick base pairing interactions<sup>27,28</sup>. These additional base pairing interactions may provide further stability to the structure, and are likely the reason for the

elevated transitional pH compared to the C-tract length. There are actually five tracts of cytosines in the sequence [5'-CCCTCCCGCCCTTGCTCCCTTCCC-3'] separated by potential loops of 1, 1, 5 and 2 nucleotides between each C-tract from the 5'- to the 3'- end. This means there is potential for more than one conformation. The presence of Watson-Crick imino protons indicates that there are some G-C base pairs (as there are no adenines in the sequence). A potential structure to accommodate these base pairs could involve a long central middle loop of eight nucleotides: 5'-CCCTCCCGCCCTTGCTCCCTTCCC-3' which would have the potential to form two CG base pairs. Interestingly, there are no TT base pairs observed. Given the presence of many thymines in the sequence and recent evidence to suggest that i-motifs are often stabilised by the presence of thymines contributing to TT base pairs in the loops<sup>29,30</sup>, this is surprising. Nevertheless, this conformational flexibility between potential different structures offers opportunity for switching by ligands, as previously demonstrated in the i-motif forming sequences from KRAS and Bcl2<sup>31,32</sup>.

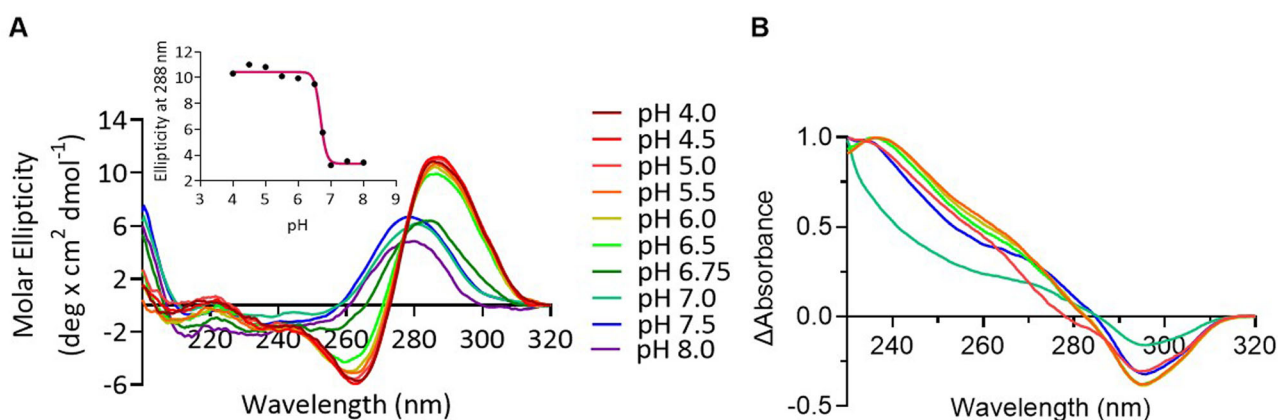
Taken together, our experiments demonstrated the Nrf2 i-motif was stable in vitro at near-neutral pH. Previous work using in-cell NMR has shown that i-motif forming sequences in cells are more stable than those in vitro<sup>16</sup>, owing in part to molecular crowding<sup>15</sup>. This indicated that the Nrf2 i-motif forming sequence in the promoter region had the potential for targeting with small molecule ligands.

### i-Motif binding ligands identified using a fluorescent indicator displacement assay

Using a previously developed thiazole orange (TO) based FID assay<sup>33</sup>, a screening experiment of 1584 drugs present in the National Cancer Institute (NCI) Developmental Therapeutics Program (DTP) was conducted to identify ligands that bind to Nrf2 at pH 6.5. We identified 11 ligands that displaced thiazole orange by greater than 20% at 2.5  $\mu$ M (Table 1) and the concentration required for 50% displacement ( $DC_{50}$ ) for each of these compounds was determined using the same method. We note that the results from compound 202386 were outside the normal expected range for error – consistent with difficulties with solubility under the conditions used in the experiments. These 11 ligands were taken forward to explore their cellular bioactivity in a cancer cell-line.

### Modulation of Nrf2 in human colon carcinoma cells

A highly drug-resistant cancer cell line, HCT-116, was selected for bioactivity screening assays due to the high expression levels of proteins involved in chemo-resistance, including Nrf2<sup>34,35</sup>. The 11 hit compounds were tested for their effect on the viability of human colon carcinoma cells following 24 h treatment at 2.5  $\mu$ M, using a cell-based MTS assay (Fig. 2A). Selecting a low concentration to identify the most potent drugs reduces the likelihood of the identification of off-target effects in downstream investigations, as well



**Fig. 1 | Biophysical Characterisation of the i-motif forming sequence from *Nrf2*.** **A** Circular dichroism spectra of Nrf2 (10  $\mu$ M DNA) in buffer containing 10 mM sodium cacodylate, 100 mM potassium chloride, at pH 4.0–8.0. Transitional pH

( $pH_T$ ) was determined using ellipticity at 288 nm (inset). **B** Thermal difference spectra of Nrf2 (2.5  $\mu$ M DNA) in buffer containing 10 mM NaCaco, 100 mM KCl, at pH 5.0, 5.5, 6.0, 6.5, 7.0 and 7.5.

Table 1 | Ligands identified from FID screening

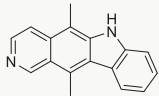
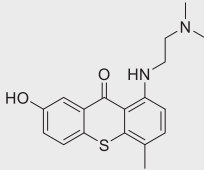
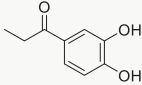
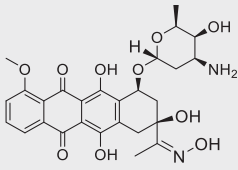
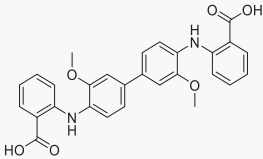
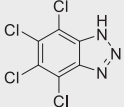
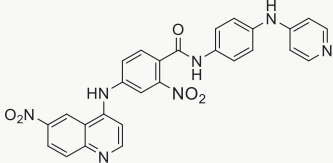
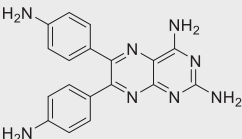
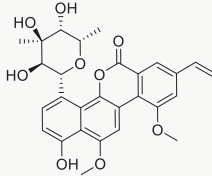
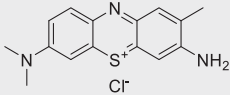
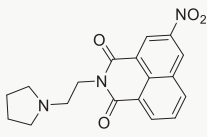
Compound	NSC	$D_{TO}$ (%)	$DC_{50}$ ( $\mu$ M)
	71795	59 ± 0.5	3.9 ± 1.0
	317003	43 ± 3.8	3.4 ± 0.5
	63842	47 ± 3.6	1.5 ± 0.1
	143491	34 ± 5.7	4.9 ± 1.5
	73735	36 ± 0.4	2.3 ± 0.1
	134674	32 ± 2.9	3.2 ± 0.1
	202386	27 ± 20.3	>20
	61642	26 ± 8.7	12 ± 3.4

Table 1 (continued) | Ligands identified from FID screening

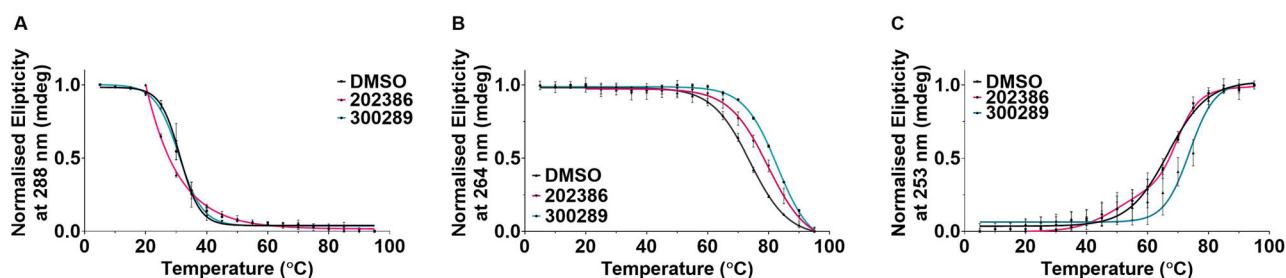
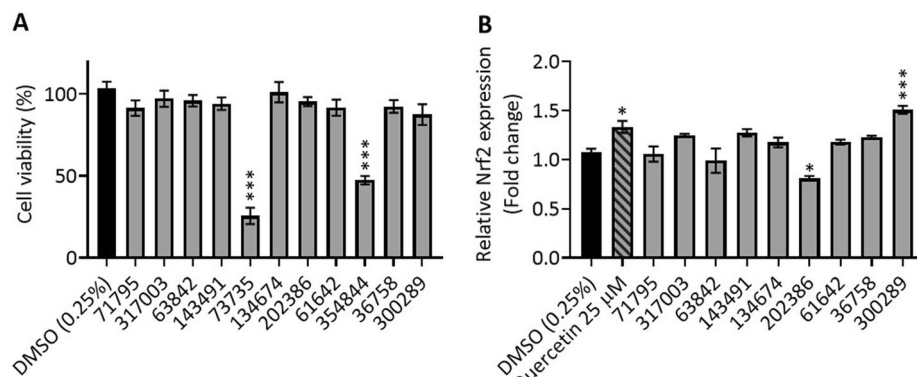
Compound	NSC	$D_{TO}$ (%)	$DC_{50}$ ( $\mu$ M)
	354844	25 ± 1.4	6.3 ± 0.6
	36758	22 ± 0.7	16 ± 2.7
	300289	20 ± 4.8	13 ± 3.0

Structures of hit compounds from the FID assay with 0.5  $\mu$ M Nrf2 in 10 mM sodium cacodylate, 100 mM potassium chloride at pH 6.5, % displacement of TO from the screening at 2.5  $\mu$ M ( $D_{TO}$ ) and the concentration required for 50% displacement of TO ( $DC_{50}$ ).

as avoiding background effects from DMSO on cell viability and *Nrf2* expression. Cell viability was reduced following treatment with compounds 73735 and 354844, and therefore excluded from gene expression experiments. The effect of the remaining ligands on *Nrf2* mRNA expression was determined 4 h after treatment, alongside quercetin as a positive control (Fig. 2B). Quercetin has been shown to bind both quadruplexes and duplex DNA<sup>36–38</sup>, and is well known to upregulate *Nrf2* in vitro<sup>39,40</sup>. As expected, quercetin at 25  $\mu$ M significantly upregulated *Nrf2* expression compared to the vehicle control (1.33-fold,  $p = 0.01$ )<sup>41</sup>. Gene expression of *Nrf2* was also found to increase in response to compound 300289 (1.51-fold,  $p \leq 0.001$ ) and was increased, though not significantly in response to 2 further ligands (compound 143491 (1.29-fold,  $p = 0.09$ ) and compound 317003 (1.25-fold,  $p = 0.17$ )). *Nrf2* expression was significantly decreased following treatment with compound 202386 (0.8-fold,  $p \leq 0.001$ ). The Nrf2 pathway is involved in drug metabolism in the cell and is upregulated in response to certain chemopreventive agents<sup>42</sup>, so it was anticipated that the presence of some of these ligands would activate this pathway. Compound 300289 is pinafide, a member of the naphthalimide family, derivatives of which have known DNA-binding and anti-tumour properties<sup>43</sup>.

Although a plethora of ligands that upregulate Nrf2 protein expression have been identified, there are relatively few that successfully inhibit its expression<sup>6</sup>. It has been demonstrated that the knockdown of *Nrf2* using miRNA in patients with acute myeloid leukaemia reduced tumour cell clonogenicity and improved their chemotherapeutic responsiveness<sup>44</sup> and it is therefore interesting that we have identified one ligand that appears to reduce *Nrf2* gene expression (compound 202386,  $p \leq 0.001$ ). Although many ligands have been shown to disrupt the interaction of Nrf2/Keap1<sup>6</sup>, this is the first example of controlling the promoter of *Nrf2* directly. The upregulation of *Nrf2* suppresses oxidative stress and inflammatory responses and increases cell survival<sup>4</sup>, which is an important mechanism of cell defence in diseases driven by low-level, chronic inflammation, including cancer<sup>45</sup>. Much focus has been given to modulating post-translational mechanisms, such as the Keap1-Nrf2 interaction and subsequent expression of antioxidant genes under *Nrf2* regulation using small molecules, such as dimethyl fumarate (DMF), which has been used for several decades for the

**Fig. 2 | Bioactivity of Nrf2 i-motif-binding ligands.** **A** Effect of Nrf2 i-motif-binding ligands on HCT-116 cell viability. Cells in monolayer were treated with 2.5  $\mu$ M of each ligand (Table 2), 0.25% DMSO, or media only for 24 h prior to the addition of MTS reagent for 4 h. Columns represent absorbance values presented as a percentage of an untreated (media only) control (mean  $\pm$  SEM,  $n = 4$ ). **B** Effect of Nrf2 i-motif binding ligands on Nrf2 mRNA expression in HCT-116 cells. Cells in monolayer were treated with 2.5  $\mu$ M of each ligand (Table 1; grey bars), 0.25% DMSO (vehicle control; black bar), quercetin (positive control; stiped bar) or media only for 4 h prior to RNA extraction, reverse transcription and RT-qPCR. Ct values were normalised to a GAPDH control and columns represent the fold change relative to an untreated control (media only) control (mean  $\pm$  SEM,  $n = 4$ ). Effects relative to vehicle control (0.25% DMSO; black bar) were determined by one-way ANOVA with post-hoc Dunnett, where  $p < 0.05$  is significant. \* $p \leq 0.05$ , \*\*\* $p \leq 0.001$ .



**Fig. 3 | DNA melting experiments in the presence of ligands.** Normalised Circular dichroism melting experiment for Nrf2C (A), Nrf2G (B), and DS (C) DNA in presence of 5 molar equivalents of DMSO control (black) NSC 202386 (pink) and

NSC 300289 (teal) in 10 mM sodium cacodylate, 100 mM potassium chloride, pH 6.5. Data are presented as Mean  $\pm$  SD,  $n = 2$ .

management of psoriasis and, more recently, relapsing-remitting multiple sclerosis<sup>1</sup>. Many more Nrf2-modulating drugs are currently in phase III testing, however, there are issues regarding specificity and off-target effects. Relatively few studies demonstrate regulation of *Nrf2* gene expression directly, so it remains to be seen whether this mechanism of action eases or exacerbates off-target effects.

### Characterisation of Nrf-2 modulating compounds by CD melting

To determine the comparative interaction between different types of DNA structures, compound 300289 and compound 202386 were subjected to further binding measurements using CD melting. We measured the melting of C-rich Nrf2C and its complementary G-rich sequence Nrf2G (5'-GGG-AAG-GGA-GCA-AGG-GCG-GGA-GGG-3'), and control double-stranded sequences (5'-GGC-ATA-GTG-CGT-GGG-CGT-TAG-C-3' plus complement). DNAs were annealed at 10  $\mu$ M in 10 mM NaCaco, 100 mM KCl at pH 6.5. CD melting experiments were performed in the presence of 5 molar equivalents of ligand (50  $\mu$ M) or the equivalent amount of DMSO without ligand as a control. Example CD melting experiments can be found in Fig. 3 and  $\Delta T_m$  values can be found in Table 2, further supporting experiments are detailed in Supplementary Fig. S3.

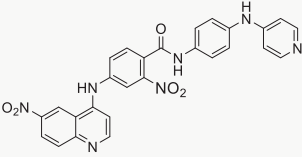
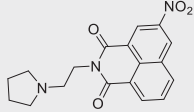
Compound 300289 was found to have no significant effect on the stability of Nrf2C ( $\Delta T_m = -1 \pm 1.0$  °C), but was found to stabilise both Nrf2G ( $\Delta T_m = +9 \pm 0.3$  °C,  $p \leq 0.001$ ) and double stranded DNA ( $\Delta T_m = +9 \pm 0.8$  °C,  $p \leq 0.001$ ). Compound 202386 was found to have a significant destabilisation effect on Nrf2C ( $\Delta T_m = -12 \pm 1.0$  °C,  $p \leq 0.01$ ), weak stabilisation of Nrf2G ( $\Delta T_m = +5 \pm 0.3$  °C,  $p \leq 0.001$ ) and no effect on

double stranded DNA ( $\Delta T_m = 0 \pm 0.7$  °C). The destabilisation effects of 202386 were complicated by poor ligand solubility at low temperatures, so measurements were only able to start from 20 °C, as opposed to from 5 °C. The quoted  $\Delta T_m$  is therefore an underestimate. It is important to note here that destabilisation potentially indicates that ligand binding stabilises a different/alternative structure to the original conformation. This could be another i-motif in the equilibrium mixture<sup>46</sup>, which is less stable than the main annealed population, an alternative partially folded species<sup>47</sup>, hairpin<sup>48</sup> or the single strand<sup>49</sup>. There are a number of compounds that destabilise i-motif structures, most of which also strongly stabilise G-quadruplex DNA<sup>11,12,49</sup>. In this case, we observe an example of a ligand that very strongly destabilises i-motif DNA, but with only weak stabilisation of G-quadruplex DNA.

### Determination of binding and specificity of Nrf-2 modulating compounds

Given the stabilising and biological properties of both 202386 and 300289, we were interested in determining their binding affinities against the i-motif and G-quadruplex forming sequences from Nrf-2. Dissociation constants ( $K_d$ s) were determined using UV titrations and are depicted in Supplementary Figs. S4–6 and summarised in Table 3. We determined the  $K_d$  of 202386 to be  $0.25 \pm 0.1$   $\mu$ M against Nrf2C (Supplementary Figs. S4A and S5A) and  $0.6 \pm 0.1$   $\mu$ M against Nrf2G (Supplementary Figs. S4B and S5B), demonstrating a two-fold preference for the Nrf-2 i-motif over the G-quadruplex. Ligand 202386 was also found to bind double helical DNA with much lower affinity ( $K_d = 1.9 \pm 0.1$   $\mu$ M, Supplementary Figs. S4A and S5G), showing a six-fold preference for the i-motif. The  $K_d$  for

**Table 2 | Stabilisation temperatures measured by CD melting**

Structure	Compound	$\Delta T_m$ (°C)		
		Nrf2C	Nrf2G	DS
	202386	-12 ± 1.0 **	+5 ± 1.0***	0 ± 0.7
	300289	-1 ± 1.0	+9 ± 0.3***	+9 ± 0.8***

10  $\mu$ M DNA annealed in 10 mM sodium cacodylate, 100 mM potassium chloride at pH 6.5. Values represent the average of two experiments and the errors are standard deviations.

202386 against Nrf2C (250 nM) is in the same range as BisA and Pyridostatin binding the i-motif forming sequence from the human telomere ( $K_d = 220$  nM and 320 nM respectively)<sup>50</sup> when measured at the same pH (6.5). Other literature ligands with similar binding affinities are PBP1 ( $K_d = 300$  nM, measured at pH 6 against Bcl2)<sup>51</sup> and compound 1 ( $K_d = 200$  nM, measured at pH 6 against Hras-1)<sup>52</sup>. To the best of our knowledge, the only ligands with tighter binding to i-motifs are IF2 and TmPyP4 ( $K_d = 34.6$  nM and 28.7 nM respectively, measured at pH 7.4 against the *BmPOUM2* i-motif)<sup>53</sup>. In contrast to the tight binding of 202386, ligand 300289 showed comparatively weaker binding to Nrf2C ( $K_d = 20 \pm 1.8$   $\mu$ M, Supplementary Figs. S4C and S6A) and Nrf2G ( $K_d = 6.6 \pm 0.9$   $\mu$ M, Supplementary Figs. S4D and S6B), with a three-fold preference for the G-quadruplex forming sequence from *Nrf2*.

To explore specificity against other i-motif and G-quadruplex forming sequences in the genome, we examined binding against sequences from the human telomere (hTeloC: 5'-TAA-CCC-TAA-CCC-TAA-CCC-TAA-CCC-3' and hTeloG: 5'-GGG-TTA-GGG-TTA-GGG-TTAG-GGT-TA-3') and the promoter region of death associated protein (DAPC: 5'-CCC-CCG-CCC-CCG-CCC-CCG-CCC-CCG-CCC-CC-3' and DAPG: 5'-GGG-GGC-GGG-GGC-GGG-GGC-GGG-GGC-GG-3'). From this, we found that 300289 binds all the G-quadruplexes with approximately similar affinity ( $K_d = 6.6 \pm 0.9$   $\mu$ M,  $6.2 \pm 0.4$   $\mu$ M and  $8.2 \pm 0.7$   $\mu$ M for Nrf2G, hTeloG and DAPG respectively, Supplementary Figs. S4D and S6B, D, E) and was more specific for the DAPC i-motif ( $K_d = 6.0 \pm 0.9$   $\mu$ M, Supplementary Figs. S4C and S6A) and hTeloC ( $K_d = 10. \pm 1.3$   $\mu$ M, Supplementary Figs. S4C and S6C) compared to Nrf2C ( $K_d = 20 \pm 1.8$   $\mu$ M). 202386 was also found to bind the other i-motifs ( $K_d = 1.1 \pm 0.2$   $\mu$ M for hTeloC and  $K_d = 0.9 \pm 0.1$   $\mu$ M for DAPC) (Supplementary Figs. S4A, S6A and C). However, 202386 was found to bind the i-motif from *Nrf2*, better than all other DNAs examined ( $K_d = 250$  nM), with between two to six-fold preference for the *Nrf2* i-motif over the other structures. This preference for binding the *Nrf2* i-motif complements its destabilisation properties.

Ligand 202386, as part of one of the NCI libraries, has previously been shown to have weak Anti-HIV-1 integrase activity<sup>54</sup>. However, previous screening against RNA-Protein interactions within the HIV-1 lifecycle<sup>55</sup> and the human telomeric G-quadruplex<sup>56</sup> did not raise this ligand as a hit. Others who have examined this compound against *Entamoeba histolytica*<sup>57</sup> and *Mycobacterium tuberculosis*<sup>58</sup> also did not find it to be a hit. Given the binding and destabilisation properties of 202386 against the i-motif forming sequence from *Nrf2* and the corresponding effects on *Nrf2* gene expression, this opens up the possibility of further compounds being developed to target *Nrf2* by binding and destabilising i-motif DNA. In this case, ligand 202386 is a promising lead for development for higher affinity analogues, with

improved general solubility and also selectivity for i-motif over other non-canonical structures. As it is available from the NCI, it also may be a useful probe for further i-motif binding studies.

## Conclusions

We have characterised a previously unexplored i-motif DNA structure in the promoter region of *Nrf2* and have determined that it is able to form at near physiological conditions. Further, we have identified a number of ligands that bind to the i-motif structure in the *Nrf2* promoter and two that can modulate *Nrf2* gene expression. Ligand 202386 was found to bind the i-motif from *Nrf2* very tightly (250 nM) with two to six-fold preference for the i-motif from *Nrf2* over other i-motifs, G-quadruplexes and double-stranded DNA. This is the first example of controlling the promoter of *Nrf2* by targeting DNA structures and offers a potential alternative mode of action for targeting this gene. The hit compounds could be used as tools for modulation of *Nrf2* or as leads for the development of novel future therapies for restoring cellular homeostasis and the efficacy of current cancer therapies.

## Methods

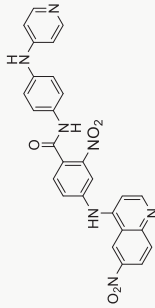
### General experimental

Oligonucleotides, purified by reverse phase HPLC, were purchased from Eurogentec. The nucleotide sequence for *Nrf2* was 5'-CCC-TCC-CGC-CCT-TGC-TCC-CTT-CCC-3' DNA stock solutions were dissolved in MilliQ water and diluted to the appropriate concentration in 10 mM sodium cacodylate, 100 mM potassium chloride at the specified pH and annealed by heating at 95 °C for 5 min and cooling to room temperature overnight. A compound library of 1584 compounds (Diversity Set IV) was obtained from the National Cancer Institute (NCI) Developmental Therapeutics Program (DTP) and was constituted to 1 mM in DMSO. Quercetin (Sigma Aldrich) was solubilised to 10 mM in DMSO and is utilised in experiments due to known duplex and quadruplex DNA binding effects and upregulation of *Nrf2*, in vitro<sup>39,40,59</sup>.

### Circular dichroism spectroscopy and melts

Nrf2C (10  $\mu$ M) was annealed in buffer at pH 4.0 to pH 8.0 to give a total volume of 200  $\mu$ L for each sample. Spectra were recorded using a Jasco J-810 spectropolarimeter with a Starna Scientific type-21 quartz cuvette of 1 mm path length. Scans were performed at 20 °C from 200 nm to 320 nm with a scanning speed of 200 nm/min, a response time of 1 s, a pitch of 0.5 nm, and a bandwidth of 2 nm. Controls containing only the respective buffer were recorded and subtracted from the experimental data. All spectra represent an average of four scans. The transitional pH (pH<sub>T</sub>) was calculated by the

Table 3 | Dissociation constants determined by UV titrations

Structure	Compound	$K_d$ ( $\mu\text{M}$ )						
		Nrf2C	hTeloC	DAPC	Nrf2G	hTeloG	DAPG	DS
	202386	0.25 ± 0.1	1.1 ± 0.2	0.9 ± 0.1	0.6 ± 0.1	1.0 ± 0.1	1.6 ± 0.2	1.9 ± 0.1
	300289	20 ± 1.8	10 ± 1.3	6.0 ± 0.9	6.6 ± 0.9	6.2 ± 0.4	8.2 ± 0.7	2.4 ± 0.2

10  $\mu\text{M}$  DNA annealed in 10 mM sodium cacodylate, 100 mM potassium chloride at pH 6.5. Values represent the average of three experiments and the errors are standard deviations.

Boltzmann sigmoidal curve fitting of the ellipticity signal at 288 nm and respective pH values.

Furthermore, CD melts were recorded on a Jasco J-1500 spectropolarimeter using a 1 mm path-length quartz cuvette. C-rich *Nrf2* and its complementary G-rich *Nrf2* (5'-GGG-AAG-GGA-GCA-AGG-GCG-G GA-GGG-3'), and control double-stranded sequences (5'-GGC-ATA-G TG-CGT-GGG-CGT-TAG-C-3' plus complement) were prepared at 10  $\mu\text{M}$  each in 10 mM sodium cacodylate, 100 mM potassium chloride at pH 6.5. CD melt was performed in the presence of 5 molar equivalents of ligand (50  $\mu\text{M}$ ) or DMSO and a spectrum from 230 nm – 320 nm at a speed of 200 nm/min was measured with increasing temperature at a rate of 1  $^{\circ}\text{C}/\text{min}$  from 5 to 95  $^{\circ}\text{C}$ . Four accumulations of scans are recorded at 5  $^{\circ}\text{C}$  intervals, 1 s response time, 0.5 nm pitch, 2 nm bandwidth. The ellipticity for Nrf2C (at 288 nm), Nrf2G (at 264 nm), and double-stranded DNA (at 253 nm) was plotted against temperature and fitted with the Boltzmann sigmoidal curve to determine the melting temperature ( $T_m$ ) at which 50% of the thermal denaturation had taken place by using Graphpad Prism (version 9).

### UV spectroscopy

All UV experiments were performed on a Jasco V760 UV-vis spectrophotometer. Thermal melting curves were derived from absorbance recordings at 295 nm. Samples were held at 4  $^{\circ}\text{C}$  for 5 min then heated to 95  $^{\circ}\text{C}$  three times at a rate of 0.25  $^{\circ}\text{C}/\text{min}$  and data was recorded every 0.1  $^{\circ}\text{C}$  during melting and annealing.  $T_M$  values were determined using first derivative methods. Thermal difference spectra were obtained by conducting a wavelength scan between 320 nm and 230 nm at 4  $^{\circ}\text{C}$  and 95  $^{\circ}\text{C}$ . The spectra of the folded structure at 4  $^{\circ}\text{C}$  was subtracted from that of the unfolded structure at 95  $^{\circ}\text{C}$ . The data were normalised and the maximum change in absorption was set to +1, as previously described<sup>25</sup>.

UV-based binding assays were determined using the wavelength at maximum absorbance of NSC 202386 at 320 nm and NSC 300289 at 335 nm in 10 mM sodium cacodylate, 100 mM potassium chloride at pH 6.5. DNA was annealed in the same buffer and added in concentrations up to 50  $\mu\text{M}$ . DNA samples used were Nrf2C [5'-CCC-TCC-CGC-CCT-TGC-TCC-CTT-CCC-3'], hTeloC [5'-TAA-CCC-TAA-CCC-TAA-CCC-TAA-CCC-3'], DAPC [5'-CCC-CCG-CCC-CCG-CCC-CCG-CCC-CCG-CCC-CC-3'] and DS [5'-GGC-ATA-GTG-CGT-GGG-CGT-TAG-C-3'] and Nrf2G [5'-GGG-TTA-GGG-TTA-GGG-TTAG-GGT-TA-3'], hTeloG [GGGTTAGGGTTAGGGTTAGGGTTA], DAPG [5'-GGG-GGC-GGG-GGC-GGG-GGC-GGG-GGC-GGC-GGC-GG-3']. Experiments were performed in triplicate. Data shown as Mean  $\pm$  SD, n = 3.

### Nuclear Magnetic Resonance (NMR)

$^1\text{H}$  NMR spectra of *Nrf2* sequence were recorded at 298 K using an 800 MHz Bruker Avance III spectrometer equipped with a triple resonance z-gradient probe<sup>10,60</sup>. To minimise saturation of labile proton signals, jump return and watergate solvent suppression pulse sequences were used. Samples were prepared to a final concentration of 10  $\mu\text{M}$  in buffer (pH 6.5). 10%  $\text{D}_2\text{O}$  was added immediately before commencing experiments. NMR spectra were acquired and processed using Bruker's Top-Spin™ software package (v3.1.7 Bruker Biospin).

### Fluorescent Indicator Displacement (FID) assay

*Nrf2*-binding ligands were identified from a library of 1584 compounds using a fluorescent indicator displacement (FID) assay<sup>33</sup>. The fluorescent indicator thiazole orange (TO) was diluted to 1  $\mu\text{M}$  in pH 6.5 buffer and the Nrf2C sequence was annealed at 50  $\mu\text{M}$  DNA concentration in pH 6.5 buffer. TO buffer was excited at 430 nm and the background fluorescence was recorded between 450 nm and 650 nm using a BMG LabTech Omega plate reader. The background fluorescence at 450 nm was normalised as 0%. Nrf2C was then added to each well to a concentration of 1  $\mu\text{M}$ , mixed, and incubated for 10 min at room temperature. Fluorescence emission intensity at 450 nm was normalised as 100%. Ligands were then added, in triplicate, to a final concentration of 2.5  $\mu\text{M}$ . To determine  $\text{DC}_{50}$  values, ligands were

titrated with spectra measured between each addition. Displacement was calculated as previously described<sup>33</sup>. Where  $D_x$  is TO displacement resulting from substitution with the candidate ligand and  $D_x = 1 - (F_x/F_0)$ ,  $F_0$  is the fluorescence intensity of the TO-i-motif complex in the absence of any ligands and  $F_x$  is the fluorescence intensity after titration with the candidate ligand.

### MTS assay

Human colon carcinoma cells (HCT-116; purchased directly from ATCC (CCL-247™), UK and were used as directed without authentication or further testing) were cultured in DMEM containing 10% FBS and 1 × penicillin/streptomycin (Gibco, UK) and plated at a cell density of 0.02 × 10<sup>6</sup> cells/well onto 96-well flat bottom plates (Greiner Bio-one, UK) 24 h prior to experiment commencement. On the day of assay, media was replaced with 100 μL fresh medium supplemented with 2.5 μM of each ligand, 0.25% DMSO (vehicle control), or media only (untreated). After 24 h, 10 μL of MTS assay reagent (CellTiter 96® AqueousOne Solution Cell Proliferation Assay; Promega, UK) was added and cells were further incubated for 4 h at 37 °C, 5% CO<sub>2</sub>, in a humidified atmosphere. Absorbance at 490 nm was read using a BMG Omega plate reader. Cell viability was determined as a percentage of absorbance against cells treated with media only (untreated). A concentration of 2.5 μM was selected based on the relatively high concentration of DMSO in the final solutions, 0.25% DMSO did not influence growth effects or cell death within 24 h.

### qRT-PCR

HCT-116 were cultured as described above and plated at a cell density of 0.2 × 10<sup>6</sup> cells/well onto 24-well plates (Greiner Bio-one, UK) 24 h prior to experiment commencement. On the day of the assay, media was replaced with 1 mL of supplemented media containing 2.5 μM ligand, 0.25% DMSO (vehicle control), 25 μM quercetin (positive control), or media only (untreated control). Cells were incubated for 4 h at 37 °C, 5% CO<sub>2</sub>, in a humidified atmosphere. Media was then removed, and the cells were lysed, RNA collected, and reverse transcribed to cDNA, using a FastLane Cell cDNA Kit (Qiagen, UK) according to the manufacturer's instructions. 25 ng cDNA from each reaction was utilised for each qRT-PCR reaction, using primers prepared and validated by PrimerDesign, UK (*Nrf2* forward primer: 5'-CCCAGCACATCCAGTCAGA-3'; reverse primer, 3'-CAGTCATCAAATACAAAGCATCT-3'), and PrecisionPlus qPCR Master Mix (without ROX). The following PCR conditions were used utilising a QIAGEN RotorGene Q-series PCR machine: activation at 95 °C for 2 min, followed by 50 cycles of DNA denaturation (10 s, 95 °C), DNA annealing and data collection (60 s, 60 °C). Ct values were normalised against a stable reference gene GAPDH (Primer Design, UK) using the 2<sup>-ΔΔCT</sup> method<sup>61</sup> and presented as the fold change against cells treated with media only (untreated). Quercetin was used as a positive control due to its known activity on *Nrf2* gene expression<sup>39</sup>.

### Reporting summary

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

### Data availability

All data generated in this study are available within the Article, Supplementary Information and can be downloaded from <https://doi.org/10.5281/zenodo.13924302>.

Received: 18 March 2024; Accepted: 4 December 2024;

Published online: 06 January 2025

### References

- Cuadrado, A. et al. Therapeutic targeting of the NRF2 and KEAP1 partnership in chronic diseases. *Nat. Rev. Drug Discov.* **18**, 295–317 (2019).
- Malhotra, D. et al. Global mapping of binding sites for Nrf2 identifies novel targets in cell survival response through ChIP-Seq profiling and network analysis. *Nucleic Acids Res.* **38**, 5718–5734 (2010).
- Kensler, T. W. & Wakabayashi, N. Nrf2: friend or foe for chemoprevention? *Carcinogenesis* **31**, 90–99 (2010).
- Jeddi, F., Soozangar, N., Sadeghi, M. R., Somi, M. H. & Samadi, N. Contradictory roles of Nrf2/Keap1 signaling pathway in cancer prevention/promotion and chemoresistance. *DNA Repair (Amst.)* **54**, 13–21 (2017).
- Rojo de la Vega, M., Chapman, E. & Zhang, D. D. NRF2 and the Hallmarks of cancer. *Cancer Cell* **34**, 21–43 (2018).
- Lv, B. et al. NRF2 inhibitors: Recent progress, future design and therapeutic potential. *Eur. J. Med. Chem.* **279**, 116822 (2024).
- Balasubramanian, S. & Neidle, S. G-quadruplex nucleic acids as therapeutic targets. *Curr. Opin. Chem. Biol.* **13**, 345–353 (2009).
- Day, H. A., Pavlou, P. & Waller, Z. A. E. i-Motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.* **22**, 4407–4418 (2014).
- Sedghi Masoud, S. & Nagasawa, K. i-Motif-binding ligands and their effects on the structure and biological functions of i-Motif. *Chem. Pharm. Bull.* **66**, 1091–1103 (2018).
- Waller, Z. A., Howell, L. A., Macdonald, C. J., O'Connell, M. A. & Searcey, M. Identification and characterisation of a G-quadruplex forming sequence in the promoter region of nuclear factor (erythroid-derived 2)-like 2 (Nrf2). *Biochem. Biophys. Res. Commun.* **447**, 128–132 (2014).
- Abdelhamid, M. A. S., Gates, A. J. & Waller, Z. A. E. Destabilization of i-Motif DNA at neutral pH by G-Quadruplex ligands. *Biochemistry* **58**, 245–249 (2019).
- Pagano, A. et al. Common G-Quadruplex binding agents found to interact with i-Motif-forming DNA: unexpected multi-target-directed compounds. *Front Chem.* **6**, 281 (2018).
- Gehring, K., Leroy, J. L. & Gueron, M. A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature* **363**, 561–565 (1993).
- 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).
- 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. (Camb.)* **46**, 1299–1301 (2010).
- Dzatzko, S. et al. Evaluation of the stability of DNA i-Motifs in the nuclei of living mammalian cells. *Angew. Chem. Int. Ed. Engl.* **57**, 2165–2169 (2018).
- Zeraati, M. et al. I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.* **10**, 631–637 (2018).
- Viskova, P. et al. In-cell NMR suggests that DNA i-motif levels are strongly depleted in living human cells. *Nat. Commun.* **15**, 1992 (2024).
- Zanin, I. et al. Genome-wide mapping of i-motifs reveals their association with transcription regulation in live human cells. *Nucleic Acids Res.* **51**, 8309–8321 (2023).
- Xie, X. et al. Recent advances in targeting the “undruggable” proteins: from drug discovery to clinical trials. *Signal Transduct. Target Ther.* **8**, 335 (2023).
- Miao, W., Hu, L., Scrivens, P. J. & Batist, G. Transcriptional regulation of NF-E2 p45-related factor (NRF2) expression by the aryl hydrocarbon receptor-xenobiotic response element signaling pathway: direct cross-talk between phase I and II drug-metabolizing enzymes. *J. Biol. Chem.* **280**, 20340–20348 (2005).
- Kypr, J., Kejnovska, I., Rencuk, D. & Vorlickova, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **37**, 1713–1725 (2009).
- Wright, E. P. et al. Epigenetic modification of cytosines fine tunes the stability of i-motif DNA. *Nucleic Acids Res.* **48**, 55–62 (2020).

24. Miglietta, G., Cogoi, S., Pedersen, E. B. & Xodo, L. E. GC-elements controlling HRAS transcription form i-motif structures unfolded by heterogeneous ribonucleoprotein particle A1. *Sci. Rep.* **5**, 18097 (2015).
25. Mergny, J. L., Li, J., Lacroix, L., Amrane, S. & Chaires, J. B. Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res.* **33**, e138 (2005).
26. Dai, J., Ambrus, A., Hurley, L. H. & Yang, D. A direct and nondestructive approach to determine the folding structure of the I-motif DNA secondary structure by NMR. *J. Am. Chem. Soc.* **131**, 6102–6104 (2009).
27. Bhattacharya, P. K., Cha, J. & Barton, J. K. 1H NMR determination of base-pair lifetimes in oligonucleotides containing single base mismatches. *Nucleic Acids Res.* **30**, 4740–4750 (2002).
28. Serrano-Chacon, I., Mir, B., Escaja, N. & Gonzalez, C. Structure of i-Motif/Duplex Junctions at Neutral pH. *J. Am. Chem. Soc.* **143**, 12919–12923 (2021).
29. Guneri, D. et al. Structural insights into i-motif DNA structures in sequences from the insulin-linked polymorphic region. *Nat. Commun.* **15**, 7119 (2024).
30. Yazdani, K. et al. Decoding complexity in biomolecular recognition of DNA i-motifs with microarrays. *Nucleic Acids Res.* **51**, 12020–12030 (2023).
31. 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).
32. 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).
33. Sheng, Q. et al. Identification of new DNA i-motif binding ligands through a fluorescent intercalator displacement assay. *Org. Biomol. Chem.* **15**, 5669–5673 (2017).
34. Guo, J. et al. Drug resistance in colorectal cancer cell lines is partially associated with aneuploidy status in light of profiling gene expression. *J. Proteome Res.* **15**, 4047–4059 (2016).
35. Ryoo, I. G., Kim, G., Choi, B. H., Lee, S. H. & Kwak, M. K. Involvement of NRF2 signaling in Doxorubicin resistance of cancer stem cell-enriched Colonospheres. *Biomol. Ther. (Seoul)* **24**, 482–488 (2016).
36. Tawani, A., Mishra, S. K. & Kumar, A. Structural insight for the recognition of G-quadruplex structure at human c-myc promoter sequence by flavonoid Quercetin. *Sci. Rep.* **7**, 3600 (2017).
37. Sun, H. et al. Spectroscopic studies of the interaction between quercetin and G-quadruplex DNA. *Bioorg. Med. Chem. Lett.* **16**, 3586–3589 (2006).
38. Bag, S., Ghosal, S., Mukherjee, M., Pramanik, G. & Bhowmik, S. Quercetin exhibits preferential binding interaction by selectively targeting HRAS1 I-Motif DNA-forming promoter sequences. *Langmuir* **40**, 10157–10170 (2024).
39. Granado-Serrano, A. B., Martin, M. A., Bravo, L., Goya, L. & Ramos, S. Quercetin modulates Nrf2 and glutathione-related defenses in HepG2 cells: Involvement of p38. *Chem. Biol. Interact.* **195**, 154–164 (2012).
40. Lee, Y. J., Lee, D. M. & Lee, S. H. Nrf2 Expression and Apoptosis in Quercetin-treated malignant Mesothelioma cells. *Mol. cells* **38**, 416–425 (2015).
41. Sun, G. Y. et al. Quercetin attenuates inflammatory responses in BV-2 microglial cells: role of MAPKs on the Nrf2 pathway and induction of Heme Oxygenase-1. *PLoS One* **10**, e0141509 (2015).
42. Kwak, M. K., Itoh, K., Yamamoto, M. & Kensler, T. W. Enhanced expression of the transcription factor Nrf2 by cancer chemopreventive agents: role of antioxidant response element-like sequences in the nrf2 promoter. *Mol. Cell Biol.* **22**, 2883–2892 (2002).
43. Nekvinda, J. et al. Synthesis of naphthalimide-carborane and metallacarborane conjugates: Anticancer activity, DNA binding ability. *Bioorg. Chem.* **94**, 103432 (2020).
44. Rushworth, S. A. et al. The high Nrf2 expression in human acute myeloid leukemia is driven by NF-kappaB and underlies its chemo-resistance. *Blood* **120**, 5188–5198 (2012).
45. Hunter, P. The inflammation theory of disease. The growing realization that chronic inflammation is crucial in many diseases opens new avenues for treatment. *EMBO Rep.* **13**, 968–970 (2012).
46. Abdelhamid, M. A. S. & Waller, Z. A. E. Tricky Topology: Persistence of folded human Telomeric i-Motif DNA at ambient temperature and neutral pH. *Front Chem.* **8**, 40 (2020).
47. Dhakal, S. et al. Coexistence of an ILPR i-motif and a partially folded structure with comparable mechanical stability revealed at the single-molecule level. *J. Am. Chem. Soc.* **132**, 8991–8997 (2010).
48. Amato, J. et al. Conformational plasticity of DNA secondary structures: probing the conversion between i-motif and hairpin species by circular dichroism and ultraviolet resonance Raman spectroscopies. *Phys. Chem. Chem. Phys.* **24**, 7028–7044 (2022).
49. Alexandrou, E., Guneri, D., Neidle, S. & Waller, Z. A. E. QN-302 demonstrates opposing effects between i-motif and G-quadruplex DNA structures in the promoter of the S100P gene. *Org. Biomol. Chem.* **22**, 55–58 (2023).
50. Bonnet, H. et al. Assessment of presumed small-molecule ligands of telomeric i-DNA by bilayer interferometry (BLI). *Chem. Commun.* **58**, 5116–5119 (2022).
51. Debnath, M. et al. Preferential targeting of i-motifs and G-quadruplexes by small molecules. *Chem. Sci.* **8**, 7448–7456 (2017).
52. Journey, S. N. et al. Probing the hras-1(Y) i-motif with small molecules. *Medchemcomm* **9**, 2000–2007 (2018).
53. Yu, G. et al. Large-scale screening of i-motif binding compounds in the silkworm, *Bombyx mori*. *Biochem. Biophys. Res. Commun.* **589**, 9–15 (2022).
54. Hong, H. et al. Identification of HIV-1 integrase inhibitors based on a four-point pharmacophore. *Antivir. Chem. Chemother.* **9**, 461–472 (1998).
55. Bell, N. M. et al. Targeting RNA-protein interactions within the human immunodeficiency virus type 1 lifecycle. *Biochemistry* **52**, 9269–9274 (2013).
56. Rahman, K. M., Tizkova, K., Reszka, A. P., Neidle, S. & Thurston, D. E. Identification of novel telomeric G-quadruplex-targeting chemical scaffolds through screening of three NCI libraries. *Bioorg. Med. Chem. Lett.* **22**, 3006–3010 (2012).
57. Saidin, S., Othman, N. & Noordin, R. In Vitro testing of potential *Entamoeba histolytica* Pyruvate Phosphate Dikinase inhibitors. *Am. J. Trop. Med. Hyg.* **97**, 1204–1213 (2017).
58. Khare, G., Kumar, P. & Tyagi, A. K. Whole-cell screening-based identification of inhibitors against the intraphagosomal survival of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **57**, 6372–6377 (2013).
59. Bhattacharjee, S., Sengupta, P. K. & Bhowmik, S. Exploring the preferential interaction of quercetin with VEGF promoter G-quadruplex DNA and construction of a pH-dependent DNA-based logic gate. *RSC Adv.* **7**, 37230–37240 (2017).
60. Abdelhamid, M. A. et al. Redox-dependent control of i-Motif DNA structure using copper cations. *Nucleic Acids Res.* **46**, 5886–5893 (2018).
61. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(Delta Delta C(T)) Method. *Methods* **25**, 402–408 (2001).

## Acknowledgements

EW was funded by a Wellcome Trust Pathfinder Grant (204515/Z/16/Z) DG was supported by a BBSRC grant (BB/W001616/1).



## Author contributions

Z.A.E.W. secured the research funding. E.F.W., D.G., and Z.A.E.W., conceived and designed the study; M.A.O. supported E.F.W. with useful insights and discussions prior to cell-based experiments. C.J.M. and E.F.W. designed, prepared, performed the NMR experiments and analysed the data. E.F.W. designed, performed and analysed the UV melting, UV difference, CD titrations, FID screening, determination of  $DC_{50}$ s, MTS and qRT-PCR experiments; D.G. designed, performed and analysed the measurements of stabilisation potential ( $\Delta T_m$ s) and binding experiments ( $K_d$ s) with 300289 and 202386. E.F.W. and D.G. prepared the figures. D.G. reviewed all the data and prepared all the data for deposition. E.F.W. wrote the first draft of the paper, D.G. and Z.A.E.W. edited and expanded the manuscript versions thereafter. All authors contributed to the review and editing of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42004-024-01387-w>.

**Correspondence** and requests for materials should be addressed to Z. A. E. Waller.

**Peer review information** Communications Chemistry thanks Carlos Gonzalez and the other, anonymous, reviewers for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025