G-quadruplexes sequester free heme in living cells

Lucas T. Gray[1,2#], Emilia Puig Lombardi[3#], Daniela Verga[4], Corinne Guetta[4], Alain Nicolas[3], Marie-Paule Teulade-Fichou[4], Arturo Londoño-Vallejo[3] and Nancy Maizels[1,5*]

Departments of [1] Biochemistry and [5] Immunology, University of Washington, 1959 NE Pacific St., Seattle, WA 98195, USA

[2] Present address: Allen Institute for Brain Science, Seattle, Washington, 98109 USA

[3] Institut Curie, Centre de Recherche, CNRS-UMR3244, PSL Research University, 75005 Paris, France

[4] Institut Curie, Centre de Recherche, CNRS-UMR9187, INSERM-U1196, PSL Research University, Sorbonne Universités, 91405 Orsay, France

[#] These authors equally contributed to this work

[*] To whom correspondence should be addressed. Tel. +1 206-221-6876; Fax +1 206- 221-6781; Email: maizels@u.washington.edu

Abstract

Heme is an essential cofactor for many enzymes, but free heme is toxic and its levels are tightly regulated. G-quadruplexes bind heme avidly *in vitro*, raising the possibility that they may sequester heme *in vivo*. If so, then treatment that displaces heme from quadruplexes is predicted to induce expression of genes involved in iron and heme homeostasis. Here we show that PhenDC3, a G-quadruplex ligand structurally unrelated to heme, displaces quadruplex-bound heme *in vitro* and alters transcription in cultured human cells, up-regulating genes that support heme degradation and iron homeostasis, and most strikingly causing a 30-fold induction of heme oxidase 1, the key enzyme in heme degradation. We propose that G-quadruplexes sequester heme to protect cells from the pathophysiological consequences of free heme. This identifies a new function for G-quadruplexes and a new mechanism for protection of cells from heme.

Keywords: G-quadruplex, heme, iron homeostasis, gene expression, oxidative damage, genomic stability

38 # Introduction

39

40 G-quadruplexes are nucleic acid structures composed of stacked 'G-quartets' [Gellert M et al.
41 1962; Sen D and Gilbert W 1988], planar arrays of four hydrogen-bonded guanines (**Figure**
42 **1A**) connected by loops composed of other nucleotides. Quadruplexes form at sequences
43 bearing the G4 motif, consisting of four or more runs of guanines punctuated by other
44 nucleotides: $G_{\geq 3}N_xG_{\geq 3}N_xG_{\geq 3}N_xG_{\geq 3}$. In the human genome, there are nearly 400,000 G4 motifs
45 allowing loops of ≤ 7 nt [Huppert JL and Balasubramanian S 2005], and more than 700,000
46 allowing loops up to 12 nt [Maizels N and Gray LT 2013]. G4 motifs are dispersed throughout
47 the genome and considerably enriched at promoters, where they flank the transcription start site
48 (TSS) and cluster at the 5' end of the first intron [Eddy J and Maizels N 2008; Eddy J and
49 Maizels N 2009]. Analyses of individual G4 motifs have identified potential regulatory roles
50 of quadruplexes in replication, transcription, splicing, and translation [Bugaut A and
51 Balasubramanian S 2012; Tarsounas M and Tijsterman M 2013; Maizels N 2015; Valton AL
52 and Prioleau MN 2016; Svikovic S and Sale JE 2016; Weldon C et al. 2017]. G4 motifs can be
53 sites of replication arrest and they are closely associated with genetic and epigenetic stability
54 [Lemmens B et al. 2015; Maizels N 2015; Svikovic S and Sale JE 2016], but G4 motifs that are
55 thermodynamically stable and therefore genetically unstable have been purged from most
56 genomes [Puig Lombardi E et al. 2019]. This makes the abundance of G4 motifs especially
57 puzzling, and suggests that they may play useful roles in cell physiology that have yet to be
58 identified.

59 G-quadruplexes bind avidly to heme *in vitro* ($K_d \approx 10$ nM range), as first shown more
60 than 20 years ago [Li Y and Sen D 1996] and subsequently pursued in many contexts (reviewed
61 in [Canale TD and Sen D 2017]). Indeed, the size, planarity and hydrophobicity of the
62 protoporphyrin IX scaffold are well suited for stacking with the terminal G-quartets while, in

63 addition, guanine or cytosine residues may provide axial coordination to the ferric ion [Li T et
64 al. 2009], altogether making heme a privileged ligand for G-quadruplexes. This binding mode
65 has been confirmed by molecular dynamics [Nasab MG et al. 2017] and docking studies [Poon
66 LC et al. 2011]. Heme is an essential cofactor in a great variety of processes involving
67 oxidation, oxygen transport and electron transport, and just as protein enzymes use heme as a
68 prosthetic group, G-rich nucleic acid aptamers can use the electron transfer capabilities of a
69 bound heme to support peroxidase reactions [Poon LC et al. 2011]. The natural affinity of
70 quadruplexes for heme suggested that quadruplexes may bind to heme *in vivo*, enabling use of
71 heme as a cofactor for nucleic acid-catalyzed reactions [Poon LC et al. 2011]. However, free
72 heme is potentially toxic even at low concentrations, as its redox-active iron can catalyze
73 formation of harmful reactive oxygen species and promoting oxidative stress [Chiabrando D et
74 al. 2014; Ponka P et al. 2017]. This raised the possibility that sequestration of heme by
75 quadruplexes might protect cells and the genome from oxidative damage by limiting free heme
76 levels while ensuring ready availability of this key cofactor.

77 Here we experimentally test the possibility that quadruplexes sequester heme in living
78 cells. While it would be very difficult to directly quantify very small amounts of intracellular
79 heme bound to quadruplexes, free heme initiates its own elimination by inducing well-
80 documented changes in gene expression, including induction of transcription of the genes
81 encoding heme oxidase 1 (HMOX1), the key enzyme in heme degradation, and ferritin (FTH1
82 and FTL1), the iron storage protein [Kappas A and Drummond GS 1986; Hamamura RS et al.
83 2007; Ghosh S et al. 2011]. Moreover, a number of ligands that interact with G-quadruplexes
84 have been shown to readily displace quadruplex-bound heme *in vitro* [Kong DM et al. 2008]
85 and are predicted to similarly displace heme bound to quadruplexes *in vivo*. We therefore
86 designed an experimental strategy that would enable us to detect release of heme from G-
87 quadruplexes by using RNA-Seq analysis to query the effect of treatment with a G-quadruplex

88 ligand predicted to displace sequestered heme on expression of genes involved in heme
89 catabolism and iron homeostasis. We chose to use the well-characterized G4 ligand PhenDC3,
90 a bisquinolinium phenanthroline derivative structurally unrelated to heme (**Figure 1A**) [De
91 Cian A et al. 2007; Piazza A et al. 2010; Halder R et al. 2012; Chung WJ et al. 2014; Perriaud
92 L et al. 2014; Lista MJ et al 2017]. PhenDC3 is an ideal compound for genome-wide studies of
93 quadruplex function *in vivo*, because it readily enters the nucleus [Lefebvre J et al. 2017]; it
94 discriminates well between quadruplex and duplex DNA; and it exhibits a very high affinity
95 (nanomolar K_d) for quadruplexes, without pronounced selectivity for any specific quadruplex
96 structure. This lack of quadruplex selectivity can be attributed to the extensive overlap of
97 PhenDC3 with the guanines in a G-quartet, in the absence of minor groove interactions
98 characteristic of some G4 ligands, which enables PhenDC3 to bind with high affinity to the
99 external G-quartets of a quadruplex [Chung WJ et al. 2014].
100 In this work, we validate displacement of G4-bound heme by PhenDC3 and confirm the
101 hypothesis that quadruplexes sequester heme *in vivo*. By RNA-seq analysis of the response of
102 human HT1080 fibrosarcoma cells to PhenDC3 treatment we show that, as predicted, PhenDC3
103 treatment significantly up-regulated pathways involved in heme metabolism and iron
104 homeostasis. Most strikingly, it caused a 30-fold induction of HMOX1, the chief enzyme in the
105 heme degradation pathway. These findings identify a new mechanism for regulation of free
106 heme levels in living cells. They also identify an unanticipated potential function for
107 quadruplexes in maintaining genomic integrity that takes advantage of the abundance of G4
108 motifs and their dispersion throughout the genome.
109

110 ## Results and Discussion

111

112 **PhenDC3 can displace quadruplex-bound hemin**

113 G-quadruplexes can associate with hemin (Fe(III)-heme) to form G4 DNAzymes which can
114 catalyze peroxidation reactions [Li Y and Sen D 1997; Travascio P et al. 1998], and
115 displacement of hemin by a G4 ligand inhibits this catalytic activity. To evaluate the ability of
116 PhenDC3 to displace hemin from a G-quadruplex, we used the chromogenic substrate 2,2'-
117 azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) ($ABTS^{2-}$) as a probe. H_2O_2-mediated
118 oxidation of $ABTS^{2-}$ produces the corresponding radical $ABTS^{\cdot-}$ which exhibits a characteristic
119 green signal readily detected by naked-eye and upon absorbance measurement. We tested
120 PhenDC3-mediated displacement of hemin from three different G-quadruplex structures, one
121 formed from an oligonucleotide representing a 22 nt G4 motif identified at the c-MYC proto-
122 oncogene [Simonsson T et al. 1998], which folds to form a parallel structure [Ambrus A et al.
123 2005]; and two others formed by oligonucleotides PS2.M (18 nt) and PS5.M (24 nt), both of
124 which have been reported to fold into antiparallel structures in the experimental conditions
125 employed and to have strong catalytic activity once bound to hemin [Travascio P et al. 1998;
126 Travascio P et al. 2001] (see *Materials and Methods*). A strong decrease of the absorption
127 spectra intensity of $ABTS^{\cdot-}$ between 500 nm and 800 nm was observed at increasing
128 concentrations of PhenDC3 (**Figure 1B**), as predicted if displacement of hemin from the
129 quadruplex structures reduces the catalytic activity of the hemin/G-quadruplex DNAzymes.
130 This trend was observed for quadruplexes formed from all three sequences tested. PhenDC3
131 seems to be able to displace hemin more efficiently from the c-MYC sequence (**Figure 1B**,
132 **left**), with catalytic activity is almost completely suppressed after the addition of 2 molar
133 equivalents (6 μM) of PhenDC3. More efficient displacement may reflect a higher affinity of
134 PhenDC3 for the c-MYC quadruplex, combined with a lower affinity of hemin towards this

135 structure, shown by the slightly reduced enzymatic activity observed in the absence of
136 PhenDC3 ($ABTS^{\cdot-}$ maximum absorption at 740 nm is lower if compared to the spectra reported
137 for PS2.M and PS5.M). The c-MYC quadruplex forms a parallel structure [Ambrus A et al.
138 2005], while the proposed structure for both PS2.M and PS5.M is antiparallel [Travascio P et
139 al. 2001]), and this may affect hemin binding. Overall, these results show that hemin is released
140 free in solution upon PhenDC3 binding to the G-quadruplex structure.

141 **Figure 1C** shows the predicted location for the lowest energy simulated structure
142 obtained upon docking of heme to a high-resolution NMR structure (PDB accession: 2MGN)
143 [Chung WJ et al. 2014] of the c-MYC promoter quadruplex (see *Materials and Methods*). The
144 heme stacks upon a loop guanine, which in turn stacks upon the G-quartets (**Figure 1C,** top).
145 PhenDC3 similarly binds with high affinity to an external G-quartet (**Figure 1C,** bottom). Thus,
146 heme and PhenDC3 are predicted to have similar favored binding sites and lie within similar
147 G-quartet to ligand distances (axial distance ~3.6 Å for heme and ~4.1 Å for PhenDC3; **Figure**
148 **1C**).

149

150 **Response to PhenDC3 correlates with G4 motif frequency**

151 To profile the transcriptome of PhenDC3-treated cells, we used RNA-seq analysis to
152 characterize gene expression in HT1080 cells, a human fibrosarcoma cell line with stable
153 diploid karyotype. Control experiments showed that treatment with 20 μM PhenDC3 for 48 hr
154 had only a modest effect on cell viability (**Figure 2A**) or cell cycle (**Figure 2B**; **Figure S1**).
155 Whole-cell RNA was isolated from cells which were untreated (DMSO) or treated with
156 PhenDC3 under these conditions, as three biological replicates (**Figure S1**). mRNA was
157 sequenced and reads were mapped to GENCODE genes in the hg19 human reference genome
158 (see *Materials and Methods*). G4 motifs were then located using a regular expression matching
159 algorithm that allowed loops of 1 to 12 nt in length, $G_{\geq 3}N_{1\text{-}12}G_{\geq 3}N_{1\text{-}12}G_{\geq 3}N_{1\text{-}12}G_{\geq 3}$,

160 N={A,T,C,G}. This search criterion identifies 722,264 G4 motifs in the human genome,
161 approximately twice as many as a search for motifs with 1 to 7 nt loops [Huppert JL and
162 Balasubramanian S 2005], and was used because the longer loop length better captures G4
163 motifs associated with biological processes [Gray LT et al. 2014].

164 PhenDC3 treatment caused significant changes in the expression profiles of HT1080
165 cells, with samples treated with 20 μM PhenDC3 or no PhenDC3 clustering uniquely upon
166 hierarchical classification (**Figure 3A**; **Figure S2A**). A total of 1,745 genes exhibited ≥2-fold
167 differential expression in response to PhenDC3 treatment, among them 1,014 up-regulated and
168 731 down-regulated genes (FDR<0.01) (**Figure 3B**; **Table S1**). Mapping G4 motif frequencies
169 of these ≥2-fold differentially expressed genes (2xDEGs) established strong correlations
170 between transcript abundance and enrichment or depletion of G4 motifs near the promoter and
171 at the 5' end of intron 1 (FDR<0.001; **Figure 3C**). This correlation was assessed by empirical
172 calculation of false discovery rates based on random sampling from all expressed genes (see
173 *Materials and Methods*). Furthermore, the proportion of genes carrying at least one G4 motif
174 in the promoter region is significantly higher among 2xDEGs than genes which were unaffected
175 (FDR>0.01) by PhenDC3 treatment (two-proportions z-test P = 0.0025 for down-regulated
176 genes, P = 0.0032 for up-regulated genes; **Figure 3D**). In particular, up-regulated genes carry
177 significantly more G4 sequences than any other gene set (**Figure 3D**).

178 We identified a very similar pattern of enrichment and depletion of G4 motifs upon
179 application of our analytical and mapping techniques to a previous array-based dataset that
180 determined the response to PhenDC3 treatment of HeLa cells [Halder R et al. 2012], a highly
181 transformed human cervical cancer cell line (**Figure S3**). These results confirm the premise that
182 PhenDC3 treatment accentuates the physiological functions of G4 motifs, and support the
183 correlation between transcript abundance and enrichment of G4 motifs near the promoter
184 (**Figure 3C**).

185 PhenDC3 appears to interact with both DNA and RNA quadruplexes to perturb gene
186 expression. For instance, positive correlations were evident with up-regulation by PhenDC3
187 and G4 motifs at regions immediately flanking the TSS (**Figure 3**), likely to reflect PhenDC3
188 interactions with DNA; on the other hand, we detect a positive correlation between differential
189 splicing events and enrichment of G4 motifs on the non-transcribed strand at the 5' end of intron
190 1 (FDR=0.004; **Figure S4**), likely reflecting PhenDC3 interactions with the G4 sequences
191 present in the pre-mRNA.
192

193 **G4 ligands induce pathways associated with heme degradation and iron homeostasis**

194 The top ten REACTOME pathways that correlated with up-regulated 2xDEGs are shown in
195 **Figure 4A** (full list is reported in **Table S2**). Two of the top pathways are directly related to
196 heme and iron homeostasis ('*Iron uptake and transport'*, '*Transferrin endocytocis and*
197 *recycling'*); and three other pathways are related via the circadian rhythm/CLOCK pathway,
198 for which the key regulator is a transcription factor (REV-ERB) that binds heme as an essential
199 cofactor [Raghuram S et al. 2007]. Gene set enrichment analysis of up-regulated 2xDEGs also
200 correlated with two other specific pathways, associated with immune responses and nerve
201 growth factor (NGF) signaling. Heme binds to proteins to modulate their function, acting as a
202 signaling molecule in a variety of biological processes, and it has been reported that these
203 include innate immunity [Figueiredo RT et al. 2007; Dutra F and Bozza MT 2014] and neuronal
204 survival [reviewed in Smith AG et al. 2011], which may contribute to these correlations. To
205 observe the interaction between the up-regulated 2xDEGs, we performed network analysis
206 which showed how these genes clustered mainly into the four afore-mentioned pathways
207 (**Figure 4A**). The network was generated by mapping the significant genes to the STRING
208 protein-protein interactome [Szklarczyk D et al. 2015] and applying a search algorithm to
209 identify first-order neighbors for each of the mapped genes (see *Materials and Methods*). We

210 generated a highly-connected first-order network (1,301 nodes and 4,038 edges), indicating
211 concerted action of the up-regulated genes in key cellular processes involving heme.

212 Particularly striking was the 30-fold (log_2FC=4.9) up-regulation of HMOX1 (**Figure**
213 **4B**; **Table S3**), which cleaves excess heme to initiate its degradation. In addition, the two chains
214 of the ferritin iron-storage complex (FTH1 and FTL) were up-regulated 5- to 7-fold. Highly
215 significant up-regulation also was evident for the mucolipin channel protein (MCOLN1), a
216 component of the cationic channel for Fe^{++} transport; a heme transporter (SLC46A1) and a
217 heme exporter (FLVCR1). PhenDC3 treatment also induced transcription of several
218 components of the vacuolar ATPase (V-ATPase) complex involved in iron and copper transport
219 (**Figure 4B**; **Table S3**).

220 The altered regulation of iron-related pathways prompted us to analyze the expression
221 of genes that encode iron-sulfur proteins. These proteins undergo iron-regulated biogenesis in
222 the mitochondrion [Lill R 2009], where G4 ligands can localize due to their lipophilic cationic
223 features and also upon binding with quadruplexes in mitochondrial DNA [Huang WC et al.
224 2015]. Of 26 genes surveyed, 3 were modestly but significantly up-regulated in response to
225 PhenDC3 treatment, and 15 were down-regulated (**Figure S6**; **Table S4**). The only gene
226 exhibiting more than 2-fold down-regulation was the G4 helicase, RTEL1 (log_2FC=-1.1). These
227 results suggest that the primary effect of PhenDC3 may be mediated by heme release, rather
228 than catabolism of heme that releases iron.

229 The HMOX1, FTH1 and FTL1 genes are all regulated by NRF2, a transcriptional
230 activator that is stabilized upon binding to heme. NRF2 competes with the BACH1 repressor
231 protein for access to antioxidant response elements (AREs), cis-acting elements that regulate
232 not only genes involved in heme and iron metabolism, but also the glutathione and thioredoxin
233 antioxidant systems, enzymes involved in detoxification of reactive oxygen species and
234 xenobiotics, and in regeneration of NADPH [reviewed in Tonelli C et al. 2018]. We therefore

235 asked how the transcriptional response of other human cells to heme or to synthetic G4 ligands
236 PhenDC3, CM03 and AQ1 (**Figure S5A**), as documented in publicly available datasets (**Figure**
237 **S5B**), compared to the response of HT1080 cells to PhenDC3. The response to PhenDC3 and
238 to heme were compared by determining the overlap of the sets of genes more than 2-fold up-
239 regulated (FDR < 0.5) in two cultured endothelial cell lines treated with Fe(III)-heme [Ghosh
240 S et al. 2011] and in HT1080 cells treated with PhenDC3. The overlap included four genes, all
241 of them targets of NRF2: HMOX1, FTH1, and two genes that function in NRF2-regulated
242 antioxidant pathways, GCLM and NQO1 (**Figure S5C**). Hierarchical cluster analysis of
243 $\log_2$(Fold Change) of these genes in all of the four publicly available datasets documenting the
244 effects of G4 ligands on gene expression supported the view that NRF2-regulated genes are key
245 targets of hemin, PhenDC3, and the anthraquinone derivative AQ1 (**Figure S5D**). In contrast,
246 a computer-designed G4 ligand of distinct structure, the trisubstituted naphthalene diimide
247 derivative, CM03 (**Figure S5A**), which is a potent inhibitor of proliferation of pancreatic tumor
248 cell lines and xenografts [Marchetti C et al. 2018], did not affect expression of these four genes.
249 These results support the hypothesis that heme is sequestered by G-quadruplexes. They
250 further show that some (but not necessarily all) G4 ligands can displace heme, and that one
251 result is activation of HMOX1 and other genes involved in heme and iron homeostasis. **Figure**
252 **5** outlines a simple model for one mechanism by which this may occur, illustrating how heme
253 displaced from a quadruplex by PhenDC3 may stabilize NRF2 to activate transcription of
254 HMOX1 and other target genes.
255 Both DNA and RNA quadruplexes bind heme *in vitro*, and either could be the source
256 of free heme released upon treatment of living cells with PhenDC3. Within a cell, interactions
257 of heme with either DNA or RNA quadruplexes could result in local oxidation that
258 compromises molecular integrity. In light of the robust catalytic activity that heme confers on
259 quadruplex DNA and RNA aptamers, Sen and colleagues proposed that accumulation of

transcripts bearing quadruplexes could be the source of potentially harmful 1- and 2-electron oxidative reactions [Poon LC et al 2011; Grigg JC et al. 2014]. Heme bound to RNA could contribute to the pathophysiology of repeat expansion diseases in which the transcript contains multimers of quadruplexes, such as the CGG repeat expansion in Fragile X; or sporadic cases of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) caused by an expansion of a GGGGCC repeat (reviewed in [Maizels N 2015]). Oxidation at heme bound to quadruplex DNA may also cause local damage, and could even constitute a potent local source of genomic instability. However, cells have evolved to repair isolated nicks and breaks in genomic DNA, and in an alternative view binding of DNA quadruplexes to heme might contribute to genomic stability, by capturing free heme near its site of origin and sequestering it until the quadruplex structure is resolved during replication or transcription. If so, sequestration of heme by quadruplexes represents an unanticipated role in maintenance of genomic stability, which may explain the abundance of G4 motifs and their dispersion throughout the genome.

Finally, the results reported here suggest that heme may function as an endogenous G4 ligand. If so, then changes in the levels of free heme in the nucleus may perturb expression not only of genes involved in heme and iron homeostasis, but also of other genes which are susceptible to perturbation by G4 ligands.

Materials and methods

Reagents for DNAzyme titration. 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) ($ABTS^{2-}$) diammonium salt, potassium acetate, DMSO, hemin and triton X-100 were purchased from Sigma-Aldrich. H_2O_2 30% was purchased from Merck. Inorganic salts and organic chemicals employed were high quality (analytical grade). Hemin and PhenDC3 were dissolved in DMSO to obtain two solutions at 1 mM concentration and stored in the dark at -20°C. $ABTS^{2-}$ was freshly prepared in water to final concentration of 100 mM. Freshly prepared H_2O_2 was made on the spot by directly diluting the high concentration H_2O_2 30% (9.8 M) to the desired concentration (2 mM) for use. The reactions were carried out in 1 x K+ buffer (40 mM HEPES-NH_4OH pH 7.8 with 20mM CH_3COOK, 1% DMSO, and 0.05% Triton X-100), optimized to favor the highest activity of G4 DNAzymes (i.e. minimal background, with the complex-catalyzed reaction occurring 20 to 50 times faster than the background reaction), prepared by dilution of 10 x K^+ buffer (400 mM HEPES-NH_4OH, pH 7.8; 200 mM CH_3COOK; 10% DMSO; 0.5% Triton x-100) made in ultrapure water and stored at 4°C before use. Assays used hemin 3 µM, $ABTS^{2-}$ 500 µM, H_2O_2 50µM, 3 µM G-quadruplexes (3 nmol); and PhenDC3 0, 0.5, 1, 3, 6 µM.

Sequences are shown in the table below for the PS2.M, PS5.M [Travascio P et al. 1998] and c-MYC [Simonsson T et al. 1998] DNA oligonucleotides used in this work. Oligonucleotides were purchased from Eurogentec (Liège, Belgique), dissolved in ultrapure water at room temperature and quantified by OD_{260nm} using the molar extinction coefficient values provided by the manufacturer.

Oligonucleotide	Sequence
PS2.M	GTGGGTAGGGCGGGTTGG
PS5.M	GTGGGTCATTGTGGGTGGGTGTGG
c-MYC	TGAGGGTGGGTAGGGTGGGTAA

302 **Hemin/G-quadruplex DNA 1:1 complex titration in the presence of increasing**
303 **concentration of PhenDC3**. Absorption titrations using the chromogenic substrate $ABTS^{2-}$
304 were performed on a Cary 300 UV-Vis spectrophotometer Agilent Technologies (Les Ulis,
305 France), using a 1 cm path-length quartz cuvette. To prepare G4 DNA, the DNA oligos (300
306 μM in 1 x K^+ buffer) were folded by heating at 95°C for 5 min and left to cool at room
307 temperature over three hr. The DNA was distributed into separate tubes to prepare the sample
308 for the titration experiments (final concentration 3 μM in 1 mL total volume for each sample).
309 One-tenth volume 10 x K^+ buffer (see above) was added to the DNA, for a final buffer
310 concentration of 1 x K^+ in ddH_2O, and the samples were left to fold over 30 min. Hemin was
311 added to a final concentration of 3 μM and the solutions were left to stand at room temperature
312 for 30 min. Samples were brought to final concentrations of 0, 0.5, 1, 3, and 6 μM PhenDC3
313 by addition of appropriate volumes of 1 mM PhenDC3 in DMSO carrier. After 30 min
314 incubation at room temperature, $ABTS^{2-}$ was then added to each sample to a final concentration
315 of 500 μM and the blanks were recorded with the spectrophotometer. To initiate the oxidation
316 reaction, H_2O_2 was added to a final concentration of 50 μM, followed by quick mixing. Five
317 min later the absorption spectra were recorded in the range of 400 to 800 nm. Final
318 concentrations were as follows: 3 μM DNA, 3 μM Hemin, 1 x K^+ (40 mM HEPES-NH_4OH,
319 pH 7.8; 20 mM CH_3COOK; 1% DMSO; 0.05% Triton x-100), 0, 0.5, 1, 3, and 6 μM PhenDC3,
320 500 μM $ABTS^{2-}$, and 50 μM H_2O_2 in a total volume of 1 mL.
321
322 **Docking of heme on the c-MYC DNA G-quadruplex.** Receptor and ligand standard
323 representation data files (Protein Data Bank, PDB format) were retrieved for the G-quadruplex
324 structure of the c-MYC promoter (PDB accession: 2MGN [Chung WJ et al. 2014]) and the
325 ligand Fe(III) protoporphyrin IX (PDB accession: HEM). We used a text editor to remove the
326 HETATM residues from the receptor PDB file, as it contains information corresponding to the

327 PhenDC3 ligand. We used the AutoDockTools (ADT) GUI v1.5.6 [Morris GM et al. 2009] to
328 prepare the pdbqt files needed for docking, which contain atomic charges and atom type
329 definitions.

330 We added hydrogen positions as well as topological information for the heme ligand to
331 perform flexible docking by allowing 10 rotatable bonds (using the *TorsionTree* utility in
332 ADT). Then, we defined a docking box covering the space to be searched around the receptor
333 (*GridBox* utility in ADT), manually adjusting the size of the box and the exact position of its
334 center before exporting it into a configuration file. We used AutoDock vina v1.1.2 [Trott O and
335 Olson AJ 2010] for MacOSX to perform docking simulations from command line, increasing
336 the exhaustiveness value from 8 (default) to 24, which gives a more consistent docking result
337 [Forli S et al. 2016]. Nine docking poses were generated at the end of the run and ranked
338 according to their binding energies, and these proved to be quite similar: all fell within the range
339 of -6.9 to -5.2 kcal/mol (mean=-6.3±0.5 kcal/mol). Root-Mean-Square Deviation of atomic
340 positions (RMSD) values were calculated relative to the best mode and are reported in the table
341 below (l.b: lower-bound; u.b: upper-bound), showing one clear preferred pose relative to the
342 receptor.

Mode	**Affinity (kcal/mol)**	**Distance from best mode**	
		RMSD l.b.	**RMSD u.b.**
1	-6.9	0	0
2	-6.8	0.655	5.719
3	-6.8	1.318	5.862
4	-6.5	1.272	1.663
5	-6.3	4.079	7.779
6	-6.0	2.242	3.312
7	-5.9	2.403	6.832
8	-5.7	4.017	9.427
9	-5.6	4.704	9.914

343

344 Visualization of the docked receptor-ligand structures was performed using PyMOL v2.3.1 for
345 MacOSX [The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC]. All

346 ligands poses were located in the same region of the G-quadruplex, from visual inspection.
347 Finally, visualization of the G4 structure bound to PhenDC3 was also performed on PyMOL,
348 using the original 2MGN PDB file.
349
350 **Cell culture and cell viability, cell death and cell cycle assays.** HT1080 fibrosarcoma cells
351 were cultured in Dulbecco's Modified Essential Medium (DMEM) with 5% fetal bovine serum,
352 glutamate, penicillin, and streptomycin. To assess cell viability, a 96 well plate was seeded with
353 2000 cells per well and cultured for 24 hr at 37°C in 5% CO_2. Samples were then brought to
354 final concentrations of 1% DMSO and 0, 10 or 20 μM PhenDC3 by addition of DMSO carrier
355 alone or 1 mM PhenDC3 in DMSO. After an additional 24 or 48 hr culture, cell viability was
356 assayed by addition of CellTiter Glo (Promega) flowed by quantification of luminescence using
357 a Fluostar Omega (BMG Labtech). Cell death was measured by staining cells with
358 LIVE/DEAD Fixable Near-IR Dead Cell Stain (Invitrogen) for 30 min, resuspending them in
359 3.7% formaldehyde, and analysis using a FACS Canto II with 633 nm excitation and 780 nm
360 emission. Cell cycle profiles of DAPI-stained cells were generated by flow cytometry on a
361 LSRII with 355 nm excitation and 450 nm emission. Results of LIVE/DEAD assays and cell
362 cycle profiles were analyzed using FlowJo v10.
363 We chose to examine cells treated with 20 μM PhenDC3 for 48 hr based on control
364 experiments which showed that effects of treatment were barely perceptible at 24 hr, and that
365 20 μM PhenDC3 was relatively non-toxic (>90% viable; **Figures 3**; **S1**). Cells were split 24 hr
366 prior to treatment and cultured in six plates at 3x105 cells per 10 cm plate, in 10 ml of media
367 per plate. Treatment was initiated by replacing culture media with 10 ml of fresh media
368 containing either 100 μl of 2 mM PhenDC3 stock in DMSO (20 μM final concentration; 3
369 culture plates), or DMSO only (3 culture plates). At 48 hr, cells were removed from each plate
370 by trypsinization and the resulting six samples separately processed for RNA-seq.

371

372 **RNA extraction and sequencing.** Cells were pelleted by centrifugation, re-suspended in
373 TRIzol Reagent (Life Technologies), and total RNA was harvested using the supplier's protocol
374 and stored at -80°C. RNA concentration was adjusted to 0.5 µg/µl, and 5 µg of each sample was
375 submitted to the UW High-Throughput Genomics Center (htSeq) for library preparation and
376 sequencing. RNA samples were tested for quality using a Bioanalyzer (Agilent Technologies).
377 mRNA library preparation was carried out using poly-dT selection. After library construction,
378 multiplexed libraries were sequenced on an Illumina Hi-Seq to 36 bp in paired-end mode. Data
379 were deposited in the Gene Expression Omnibus (GEO) database, accession number
380 GSE60630.

381

382 **RNA-seq data analysis.** Reads were aligned to the human reference genome hg19 using STAR
383 v2.7 [Dobin A et al. 2013] with the ENCODE processing pipeline standard settings for RNA-
384 seq data (*--outFilterType* BySJout, *--outFilterMultimapNmax* 20, *--alignSJoverhangMin* 8, --
385 *alignSJDBoverhangMin* 1, *--outFilterMismatchNmax* 999, *--alignIntronMin* 20, --
386 *alignIntronMax* 1000000 and *--alignMatesGapMax* 1000000). Reads mapping uniquely to
387 GENCODE-annotated genes were summarised using *featureCounts* [Liao Y et al. 2014].

388 The raw gene count matrix was imported into the R environment [R Core Team 2018]
389 for further processing and analysis. Genes with low read counts (less than ~10 reads in more
390 than 3 samples) were filtered out, leaving a set of 20,000 genes to test for differential expression
391 between control (DMSO) and treated (+PhenDC3) conditions. Differential expression analysis
392 was carried out using the R package *DESeq2* version 1.18.1 [Love MI et al. 2014]. Unless
393 otherwise stated, differentially expressed genes were identified based on false discovery rate
394 (FDR, using Benjamini-Hochberg adjusted *p*-values), with the threshold FDR < 1%. A fold-
395 change cutoff was used to create the sets of genes used for G4 enrichment analysis: only ≥2-

396 fold (≥1-log_2fold) differentially expressed genes (2xDEGs) were retained. Hierarchical
397 clustering was performed for the top 500 differentially expressed genes using Euclidean
398 distances. The *cuffdiff* tool from Cufflinks v2.1.1 [Trapnell C et al. 2013] was used to calculate
399 differentially expressed isoforms with the default q-value cutoff of < 0.05. Statistics for isoform
400 expression levels and differential isoform expression are available in the GEO repository
401 (GSE60630).
402
403 **G4 motif enrichment analysis.** G4 motifs were defined as carrying 4 or more runs of 3 or more
404 Gs, separated by 1-12 nt loops, and were located using regular expression matching as
405 previously described [Gray LT et al. 2014]. Results were stored as bed files with coordinates
406 as well as wig files for visualization and further processing. G4 enrichment was assessed using
407 R scripts that calculated the frequency of G4 overlap at each position and on each strand near
408 the TSS (+/- 250 bp) and near the 5' end of first intron (1-250 bp). TSS and intron locations
409 were retrieved from the Ensembl genes table in the UCSC hg19 iGenomes database using gene
410 symbols. Briefly, we determined false discovery rates (FDR) by selecting 1000 sets of genes of
411 matching size (e.g. for the 716 genes up-regulated >2-fold, we assessed 1000 sets of 716
412 random genes) from the pool of all reliably detected genes, which were subjected to the same
413 count analysis described above. G4 motif enrichment or depletion found in fewer than 5% of
414 the randomly selected sets was considered significant (FDR < 0.05). R scripts were used to
415 tabulate and assemble results.
416
417 **Gene Set Enrichment Analysis (GSEA) and protein-protein interaction network.** The
418 REACTOME pathways enrichment analysis was done using the GSEA (Gene Set Enrichment
419 Analysis) software [Subramanian A et al. 2005] with the Molecular Signatures Database
420 collection on the identified up-regulated gene lists. Significantly enriched pathways were

421 selected based on FDR q-values (*P*-value adjustment for multiple hypergeometric tests), with
422 the cutoff FDR<5%. The NetworkAnalyst platform [Xia J et al. 2015] and the Cytoscape 3.5
423 software [Shannon P et al. 2003] were used to create and visualize the first-order (proteins that
424 directly interact with a given protein) protein-protein interactions based on the high-confidence
425 STRING interactome information [Szklarczyk D et al. 2015]. High-confidence interactions
426 were defined has having a confidence score > 0.9 and experimental evidence. As the first-order
427 network was very dense, we constructed a minimally connected network containing all the seed
428 genes as well as essential non-seed interactors that maintain the network connection. To do this,
429 we compute pair-wise shortest paths between all seed nodes, and remove the nodes that are not
430 on the shortest paths. The obtained minimum first-order network consisted of 1,301 nodes and
431 4,038 edges, with larger nodes indicating higher connectivity.
432

Acknowledgments

Funding. This work was supported by the US National Institutes of Health National Cancer Institute (P01CA077852 to N.M.), by the Agence Nationale de la Recherche (ANR 14-CE35-0003-02 to A.N. and M.P.-T.F.) and by the PIC3i program from Institut Curie (n°91730 "Prospects of Anticancer" to A.N., M.P.-T.F. and A.L.V.). E.P.L. is a recipient of a doctoral fellowship from the French Ministry of Education, Research and Technology.

Authors' contributions. L.T.G. carried out cell culture, PhenDC3 treatment and prepared RNA. L.T.G. and E.P.L. designed the *in silico* analytical framework and wrote the code for analyzing and presenting the data. E.P.L. performed the final version of the RNA-seq data analysis and implemented the docking simulations. D.V performed the hemin/G-quadruplex DNAzyme assay. C.G. synthesized PhenDC3. A.N., M.P.-T.F., A.L.V. and N.M. conceived of and supervised the project. The manuscript was written by L.T.G., E.P.L. and N.M. All authors read and approved the final draft of the manuscript.

References

[1] Gellert M, Lipsett MN, Davies DR. (1962) Helix formation by guanylic acid. *Proc Natl Acad Sci USA*, **48**:2014-2018.

[2] Sen D, Gilbert W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*, **334**: 364–366.

[3] Huppert JL, Balasubramanian S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res*, **33**:2908-2916.

[4] Maizels N, Gray LT. (2013) The G4 genome. *PLoS Genet*, **9**:e1003468.

457 [5] Eddy J, Maizels N. (2008) Conserved elements with potential to form polymorphic G-
458 quadruplex structures in the first intron of human genes. *Nucleic Acids Res*, **36**:1321-
459 1333.
460 [6] Eddy J, Maizels N. (2009) Selection for the G4 DNA motif at the 5' end of human genes.
461 *Mol Carcinog*, **48**:319-325.
462 [7] Bugaut A, Balasubramanian S. (2012) 5'-UTR RNA G-quadruplexes: translation
463 regulation and targeting. *Nucleic Acids Res*, **40**:4727-4741.
464 [8] Tarsounas M, Tijsterman M. (2013) Genomes and G-quadruplexes: for better or for
465 worse. *J Mol Biol*, **425**:4782-4789.
466 [9] Maizels N. (2015) G4-associated human diseases. *EMBO Rep*, **16**:910-922.
467 [10] Valton AL, Prioleau MN. (2016) G-Quadruplexes in DNA Replication: A Problem or a
468 Necessity? *Trends Genet*, **32**:697-706.
469 [11] Svikovic S, Sale JE. (2016) The Effects of Replication Stress on S Phase Histone
470 Management and Epigenetic Memory. *J Mol Biol*, **429**(13):2011-2029.
471 [12] Weldon C, Behm-Ansmant I, Hurley LH, Burley GA, Branlant C, Eperon IC,
472 Dominguez C. (2017) Identification of G-quadruplexes in long functional RNAs using
473 7-deazaguanine RNA. *Nat Chem Biol*, **13**:18-20.
474 [13] Lemmens B, van Schendel R, Tijsterman M. (2015) Mutagenic consequences of a single
475 G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers. *Nat*
476 *Commun*, **13**(6): 8909.
477 [14] Puig Lombardi E, Holmes A, Verga D, Teulade-Fichou MP, Nicolas A, Londoño-
478 Vallejo A. (2019) Thermodynamically stable and genetically unstable G-quadruplexes
479 are depleted in genomes across species. *Nucleic Acids Res*, pii: gkz463 [Epub ahead of
480 print].

481 [15] Li Y, Sen D. (1996) A catalytic DNA for porphyrin metallation. *Nat Struct Biol*, **3**:743-
482 747.
483 [16] Canale TD, Sen D (2017) Hemin-utilizing G-quadruplex DNAzymes are strongly active
484 in organic co-solvents. *Biochim Biophys Acta*, **1861**:1455-1462.
485 [17] Li T, Dong S, Wang E. (2009) G-quadruplex aptamers with peroxidase-like DNAzyme
486 functions: which is the best and how does it work? *Chem Asian J*, **4**:918-22.
487 [18] Nasab GM, Hassani L, Mohammadi Nejad S, Norouzi D. (2017) Interaction of hemin
488 with quadruplex DNA. *J Biol Phys*, **43**:5-14.
489 [19] Poon LC, Methot SP, Morabi-Pazooki W, Pio F, Bennet AJ, Sen D. (2011) Guanine-
490 rich RNAs and DNAs that bind heme robustly catalyze oxygen transfer reactions. *J Am*
491 *Chem Soc*, **133**:1877-1884.
492 [20] Chiabrando D, Vinchi F, Fiorito V, Mercurio S, Tolosano E. (2014) Heme in
493 pathophysiology: a matter of scavenging, metabolism and trafficking across cell
494 membranes. *Front Pharmacol*, **5**:61.
495 [21] Ponka P, Sheftel AD, English AM, Scott Bohle D, Garcia-Santos D. (2017) Do
496 Mammalian Cells Really Need to Export and Import Heme? *Trends Biochem Sci*,
497 **42**:395-406.
498 [22] Kappas A, Drummond GS. (1986) Control of heme metabolism with synthetic
499 metalloporphyrins. *J Clin Invest*, **77**(2):335-339.
500 [23] Hamamura RS, Ohyashiki JH, Kurashina R, Kobayashi C, Zhang Y, Takaku T,
501 Ohyashiki K. (2007) Induction of heme oxygenase-1 by cobalt protoporphyrin enhances
502 the antitumour effect of bortezomib in adult T-cell leukaemia cells. *Br J Cancer*,
503 **97**(8):1099-105.

504 [24] Ghosh S, Tan F, Yu T, Li Y, Adisa O, Mosunjac M, Ofori-Acquah SF. (2011) Global
505 gene expression profiling of endothelium exposed to heme reveals an organ-specific
506 induction of cytoprotective enzymes in sickle cell disease. *PLoS One*, **6**(3):e18399.
507 [25] Kong DM, Wu J, Ma YE, Shen HX. (2008) A new method for the study of G-quadruplex
508 ligands. *Analyst*, **133**:1158-1160.
509 [26] De Cian A, Delemos E, Mergny JL, Teulade-Fichou MP, Monchaud D. (2007) Highly
510 efficient G-quadruplex recognition by bisquinolinium compounds. *J Am Chem Soc*,
511 **129**:1856-1857.
512 [27] Piazza A, Boule JB, Lopes J, Mingo K, Largy E, Teulade-Fichou MP, Nicolas A. (2010)
513 Genetic instability triggered by G-quadruplex interacting Phen-DC compounds in
514 Saccharomyces cerevisiae. *Nucleic Acids Res*, **38**:4337-4348.
515 [28] Halder R, Riou JF, Teulade-Fichou MP, Frickey T, Hartig JS. (2012) Bisquinolinium
516 compounds induce quadruplex-specific transcriptome changes in HeLa S3 cell lines.
517 *BMC Res Notes*, **5**:138.
518 [29] Chung WJ, Heddi B, Hamon F, Teulade-Fichou MP, Phan AT. (2014) Solution structure
519 of a G-quadruplex bound to the bisquinolinium compound Phen-DC3. *Angew Chem Int*
520 *Ed Engl*, **53**:999-1002.
521 [30] Perriaud L, Marcel V, Sagne C, Favaudon V, Guedin A, De Rache A, Guetta C, Hamon
522 F, Teulade-Fichou MP, Hainaut P et al. (2014) Impact of G-quadruplex structures and
523 intronic polymorphisms rs17878362 and rs1642785 on basal and ionizing radiation-
524 induced expression of alternative p53 transcripts. *Carcinogenesis*, **35**:2706-2715.
525 [31] Lista MJ, Martins RP, Billant O, Contesse MA, Findakly S, Pochard P, Daskalogianni
526 C, Beauvineau C, Guetta C, Jamin C et al. (2017) Nucleolin directly mediates Epstein-
527 Barr virus immune evasion through binding to G-quadruplexes of EBNA1 mRNA. *Nat*
528 *Commun*, **8**:16043.

529 [32] Lefebvre J, Guetta C, Poyer F, Mahuteau-Betzer F, Teulade-Fichou MP. (2017) Copper-
530 Alkyne Complexation Responsible for the Nucleolar Localization of Quadruplex
531 Nucleic Acid Drugs Labeled by Click Reactions. *Angew Chem Int Ed Engl*, **56**:11365-
532 11369.
533 [33] Li Y, Sen D. (1997) Toward an efficient DNAzyme. *Biochemistry*, **36**(18):5589-99.
534 [34] Travascio P, Li Y, Sen D. (1998) DNA-enhanced peroxidase activity of a DNA-
535 aptamer-hemin complex. *Chem Biol*, **5**(9):505-17.
536 [35] Simonsson T, Pecinka P, Kubista M. (1998) DNA tetraplex formation in the control
537 region of c-myc. *Nucleic Acids Res*, **26**(5):1167-72.
538 [36] Ambrus A, Chen D, Dai J, Jones RA, Yang D. (2005) Solution structure of the
539 biologically relevant G-quadruplex element in the human c-MYC promoter.
540 Implications for G-quadruplex stabilization. *Biochemistry*, **44**(6):2048-58.
541 [37] Travascio P, Witting PK, Mauk AG, Sen D. (2001) The peroxidase activity of a hemin-
542 -DNA oligonucleotide complex: free radical damage to specific guanine bases of the
543 DNA. *J Am Chem Soc*,**123**(7):1337-48.
544 [38] Gray LT, Vallur AC, Eddy J, Maizels N. (2014) G-quadruplexes are genome-wide
545 targets of transcriptional helicases XPB and XPD. *Nat Chem Biol*, **10**:313-318.
546 [39] Raghuram S, Stayrook KR, Huang P, Rogers PM, Nosie AK, McClure DB, Burris LL,
547 Khorasanizadeh S, Burris TP, Rastinejad F. (2007) Identification of heme as the ligand
548 for the orphan nuclear receptors REV-ERBalpha and REV-ERBbeta. *Nat Struct Mol*
549 *Biol*, **14**:1207-1213.
550 [40] Figueiredo RT, Fernandez PL, Mourao-Sa DS, Porto BN, Dutra FF, Alves LS, Oliveira
551 MF, Oliveira PL, Graça-Souza AV, Bozza MT. (2007) Characterization of heme as
552 activator of Toll-like receptor 4. *J Biol Chem*, **282**(28):20221-9.

553 [41] Dutra F and Bozza MT. (2014) Heme on innate immunity and inflammation. *Front*
554 *Pharmacol*, **5**:115.
555 [42] Smith AG, Raven EL, Chernova T. (2011) The regulatory role of heme in neurons.
556 *Metallomics*, **3**(10):955-62.
557 [43] Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J,
558 Simonovic M, Roth A, Santos A, Tsafou KP et al. (2015) STRING v10: protein-protein
559 interaction networks, integrated over the tree of life. *Nucleic Acids Res*, **43**(Database
560 issue):D447-52.
561 [44] Lill R. (2009) Function and biogenesis of iron-sulphur proteins. *Nature*, **460**:831-838.
562 [45] Tonelli C, Chio IIC, Tuveson DA. (2018) Transcriptional Regulation by Nrf2. *Antioxid*
563 *Redox Signal*, **29**(17):1727-1745.
564 [46] Huang WC, Tseng TY, Chen YT, Chang CC, Wang ZF, Wang CL, Hsu TN, Li PT,
565 Chen CT, Lin JJ et al. (2015) Direct evidence of mitochondrial G-quadruplex DNA by
566 using fluorescent anti-cancer agents. *Nucleic Acids Res*, **43**:10102-10113.
567 [47] Marchetti C, Zyner KG, Ohnmacht SA, Robson M, Haider SM, Morton JP, Marsico G,
568 Vo T, Laughlin-Toth S, Ahmed AA et al. (2018) Targeting Multiple Effector Pathways
569 in Pancreatic Ductal Adenocarcinoma with a G-Quadruplex-Binding Small Molecule.
570 *J Med Chem*, **61**(6):2500-2517.
571 [48] Zorzan E, Elgendy R, Giantin M, Dacasto M, Sissi C. (2018) Whole-Transcriptome
572 Profiling of Canine and Human in Vitro Models Exposed to a G-Quadruplex Binding
573 Small Molecule. *Sci Rep*, **8**(1):17107.
574 [49] Grigg JC, Shumayrikh N, Sen D. (2014) G-quadruplex structures formed by expanded
575 hexanucleotide repeat RNA and DNA from the neurodegenerative disease-linked
576 C9orf72 gene efficiently sequester and activate heme. *PLoS One*, **9**:e106449.

577 [50] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ.
578 (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor
579 flexiblity. *J Computational Chemistry*, **6**:2785-91.
580 [51] Trott O, Olson AJ. (2010) AutoDock Vina: improving the speed and accuracy of
581 docking with a new scoring function, efficient optimization and multithreading. *J*
582 *Computational Chemistry*, **31**:455-461.
583 [52] Forli S, Huey R, Pique ME, Sanner MF, Goodsell DS, Olson AJ. (2016) Computational
584 protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat Protoc*,
585 **11**(5):905-19.
586 [53] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,
587 Gingeras TR. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*,
588 **29**(1):15–21.
589 [54] Liao Y, Smyth GK, Shi W. (2014) featureCounts: an efficient general purpose program
590 for assigning sequence reads to genomic features. *Bioinformatics*, **30**(7):923-30.
591 [55] R Core Team (2018) R: A language and environment for statistical computing. R
592 Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
593 [56] Love MI, Huber W, Anders S. (2014) Moderated estimation of fold change and
594 dispersion for RNA-seq data with DESeq2. *Genome Biol*, **15**:550.
595 [57] Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. (2013)
596 Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat*
597 *Biotechnol*, **31**:46-53.
598 [58] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA,
599 Paulovich A, Pomeroy SL, Golub TR, Lander ES. et al. (2005) Gene set enrichment
600 analysis: a knowledge-based approach for interpreting genome-wide expression
601 profiles. *Proc Natl Acad Sci U S A*, **102**:15545-15550.

602 [59] Xia J, Gill EE, Hancock RE. (2015) NetworkAnalyst for statistical, visual and network-
603 based meta-analysis of gene expression data. *Nat Protoc*, **10**(6):823-44.
604 [60] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N,
605 Schwikowski B, Ideker T. (2003) Cytoscape: a software environment for integrated
606 models of biomolecular interaction networks. *Genome Res*, **13**(11):2498-504.

607 ## Figure legends

608

609 **Figure 1. PhenDC3 displaces hemin from three different G-quadruplex structures.**

610 **(A)** Structures of a G-quartet, Fe(III)-heme and PhenDC3 (structures are not size-scaled). **(B)**

611 Absorption spectra analyzing H_2O_2-dependent oxidation of $ABTS^{2-}$ to $ABTS^{1-}$ by G4

612 DNAzyme in the presence of the G4 ligand PhenDC3 (0-6 μM). G4 DNAzymes assayed were

613 complexes of hemin with quadruplexes formed from the indicated oligonucleotides: hemin/c-

614 MYC, hemin/PS2.M and hemin/PS5.M. **(C)** Above: structure of Fe(III) hemin (PDB accession:

615 HEM) docked on the c-MYC promoter DNA G-quadruplex, shown as (from left to right) a side

616 view, a zoomed side view and a top view. The distance from the loop to the iron at the center

617 of the heme is ~3.6 Å (middle panel, dotted yellow line). Below: model view of the solution

618 structure of the G4 ligand PhenDC3 bound to the c-MYC G-quadruplex (PDB coordinates

619 accession: 2MGN), shown as (from left to right) a side view, a zoomed side view and a top

620 view. For the PyMOL visualizations, we used default atom colors. In particular, hemin iron is

621 dark orange (RGB code triples R=0.87843137; G=0.400000000 and B=0.200000000); refer to

622 https://pymolwiki.org/index.php/Color_Values).

623

624 **Figure 2. Effects of PhenDC3 on cell viability and cell cycle.**

625 **(A)** Viability of cells treated with DMSO carrier alone (1%) or 10 or 20 μM PhenDC3 in DMSO

626 carrier as measured by ATP concentration (left) and live/dead staining (right). Error bars

627 indicate standard deviation among triplicate samples. *, $p < 0.01$; **, $p < 0.001$; ***, $p < 0.0001$.

628 **(B)** DAPI cell cycle profiles for populations treated as in panel A for 48 hr.

629

630

631 **Figure 3. PhenDC3 treatment induces significant transcriptional changes in HT1080**
632 **fibrosarcoma cells.**
633 **(A)** Supervised clustering of log-transformed gene counts obtained from three independent
634 RNA-seq experiments performed in HT1080 cells treated with PhenDC3 (20 μM) for 48 hours.
635 Top scoring 500 genes are shown. **(B)** Volcano plot of genes differentially expressed (FDR <
636 0.01) in cells treated or untreated with PhenDC3 (20 μM). A subset of the genes significantly
637 upregulated (Log_2(Fold Change) > 0) is shown. **(C)** G4 ($[G_{3+}N_{1-12}]_{3+}G_{3+}$) enrichment and
638 depletion statistics for 250 nt windows flanking the TSS and downstream of the 5' end of first
639 introns. Enrichment (green) and depletion (red) are relative to 1000 sets of randomly selected
640 gene drawn from a pool of all genes expressed at detectable levels. FDR values in parentheses.
641 FDR > 5% (0.05) is considered non-significant (n.s, no highlight). **(D)** G4 ($[G_{3+}N_{1-12}]_{3+}G_{3+}$)
642 detection in promoter regions (1000 nt upstream and 250 nt downstream of TSS) of 2xDEGs
643 (red, up-regulated; blue, down-regulated genes) and unaltered genes (gray, FDR>0.01). Pie
644 charts show the number of genes carrying at least one G4 motif (black) or no G4 motif (gray)
645 within promoter regions. Boxplots: ***, pairwise Wilcoxon tests adj$P$ < 0.001 compared to NS;
646 Pie charts: **, two-proportion z-tests P < 0.01.
647
648 **Figure 4. PhenDC3 treatment activates genes involved in heme and iron homeostasis.**
649 **(A)** Minimal first-order protein-protein interaction networks between up-regulated 2xDEGs.
650 Node size is proportional to the number of its connections with other nodes (n = 1,301 nodes).
651 The barplot shows REACTOME pathway-based enrichment analysis of genes strongly up-
652 regulated (Log_2(Fold Change) > 1; FDR < 0.01) in treated (+PhenDC3) relative to untreated
653 samples. The top 10 results correlated with upregulated genes are shown. FDR, GSEA false
654 discovery rate q-value. **(B)** Log_2(Fold Changes) for common genes involved in iron
655 homeostasis REACTOME pathways.

656

657 **Figure 5. One possible mechanism for induction of HMOX1 expression in response to**

658 **heme displacement by a quadruplex ligand.**

659 Heme sequestered by a G-quadruplex is displaced by PhenDC3. The free heme thus released

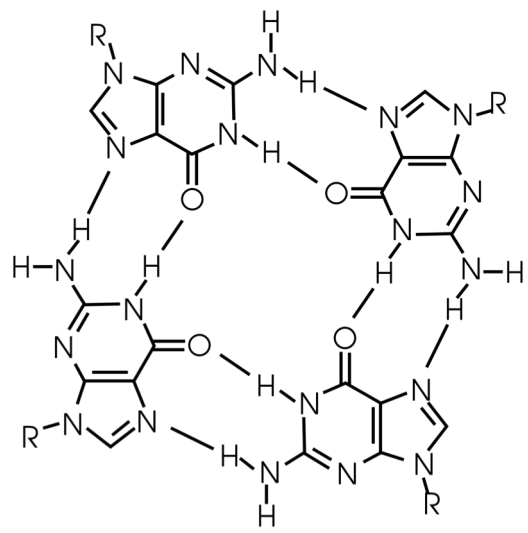
660 binds to and stabilizes NRF2, which then drives the BACH1 repressor from the promoter to

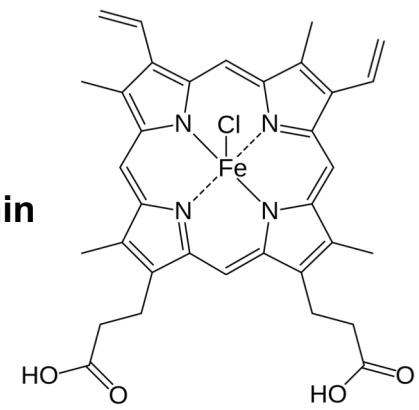
661 induce gene expression. Regulation of HMOX1 is shown, but this model could be applicable

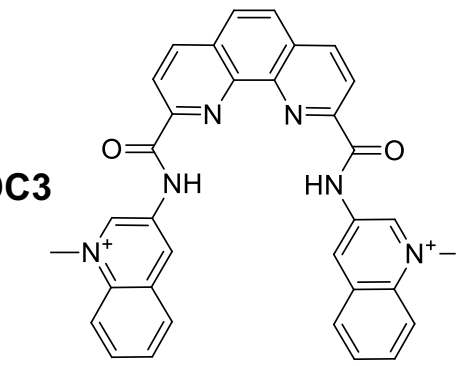
662 to many NRF2 target genes. ARE, Antioxidant Response Element.

Figure 1

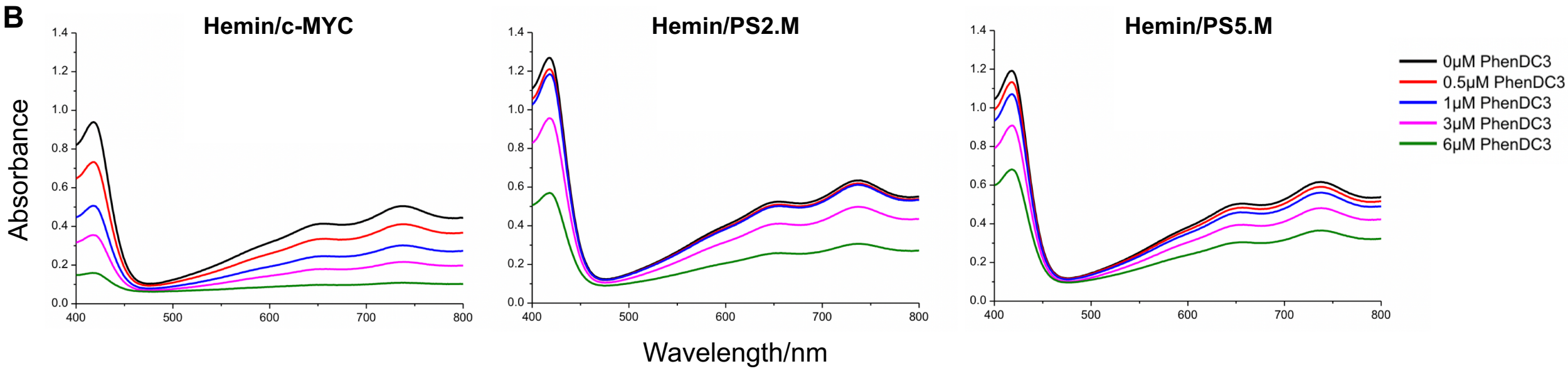
A


G-quartet


Hemin


PhenDC3

B


Hemin/c-MYC
Hemin/PS2.M
Hemin/PS5.M
Absorbance
Wavelength/nm
0μM PhenDC3
0.5μM PhenDC3
1μM PhenDC3
3μM PhenDC3
6μM PhenDC3

C

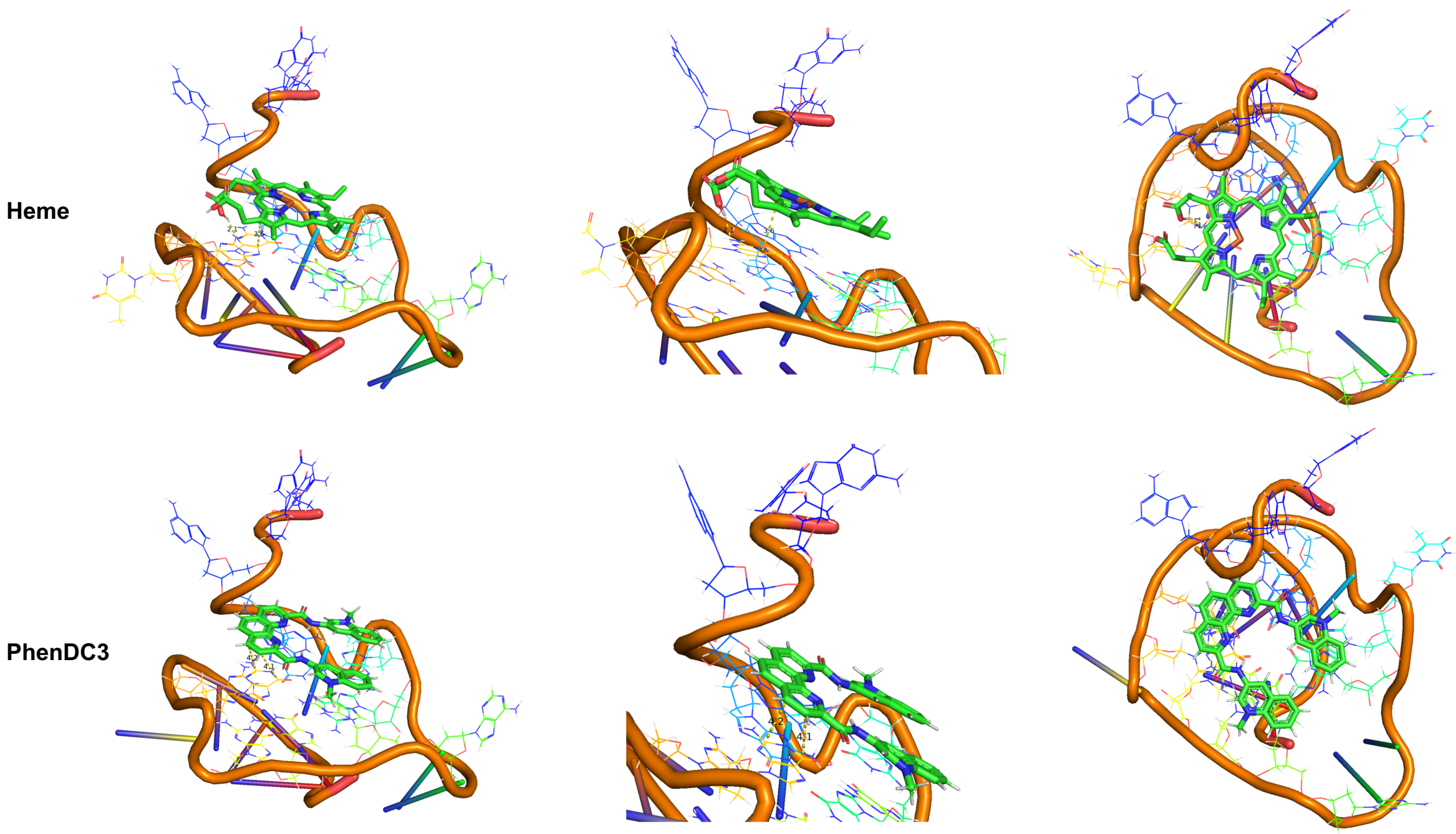
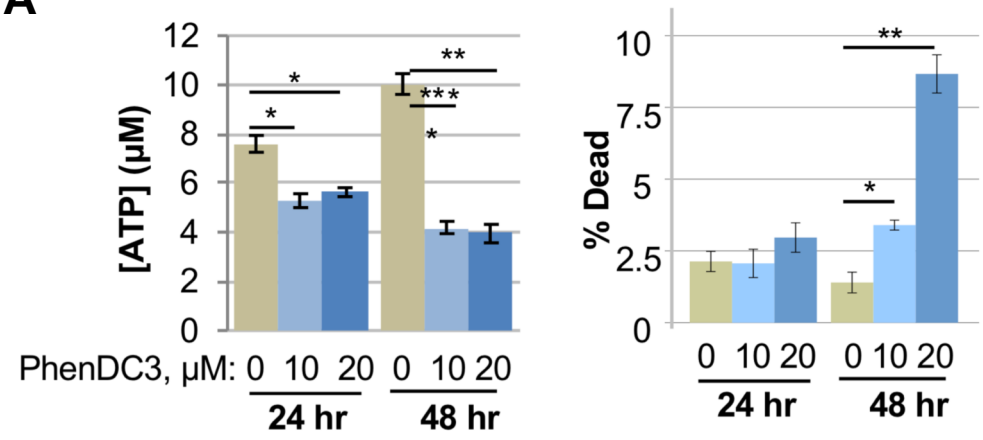

Heme
PhenDC3

Figure 2

A


[ATP] (μM)
12
10
8
6
4
2
0
*
*
**
***
*
PhenDC3, μM: 0 10 20 0 10 20
24 hr
48 hr
% Dead
10
7.5
5
2.5
0
**
*
0 10 20 0 10 20
24 hr
48 hr

B

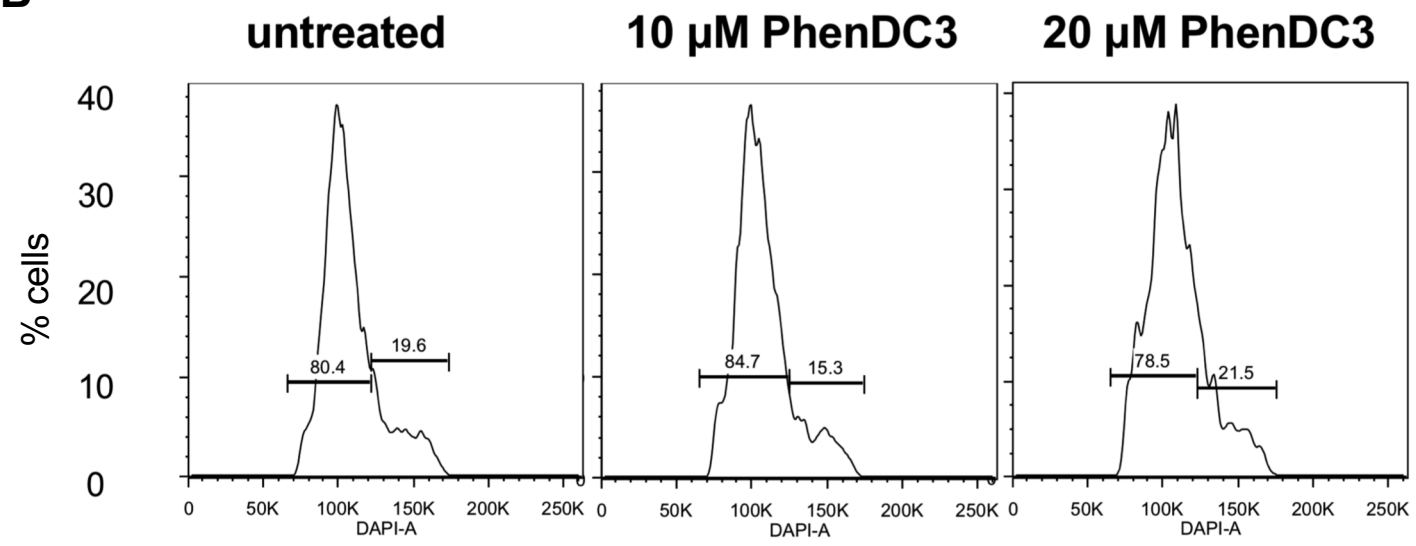
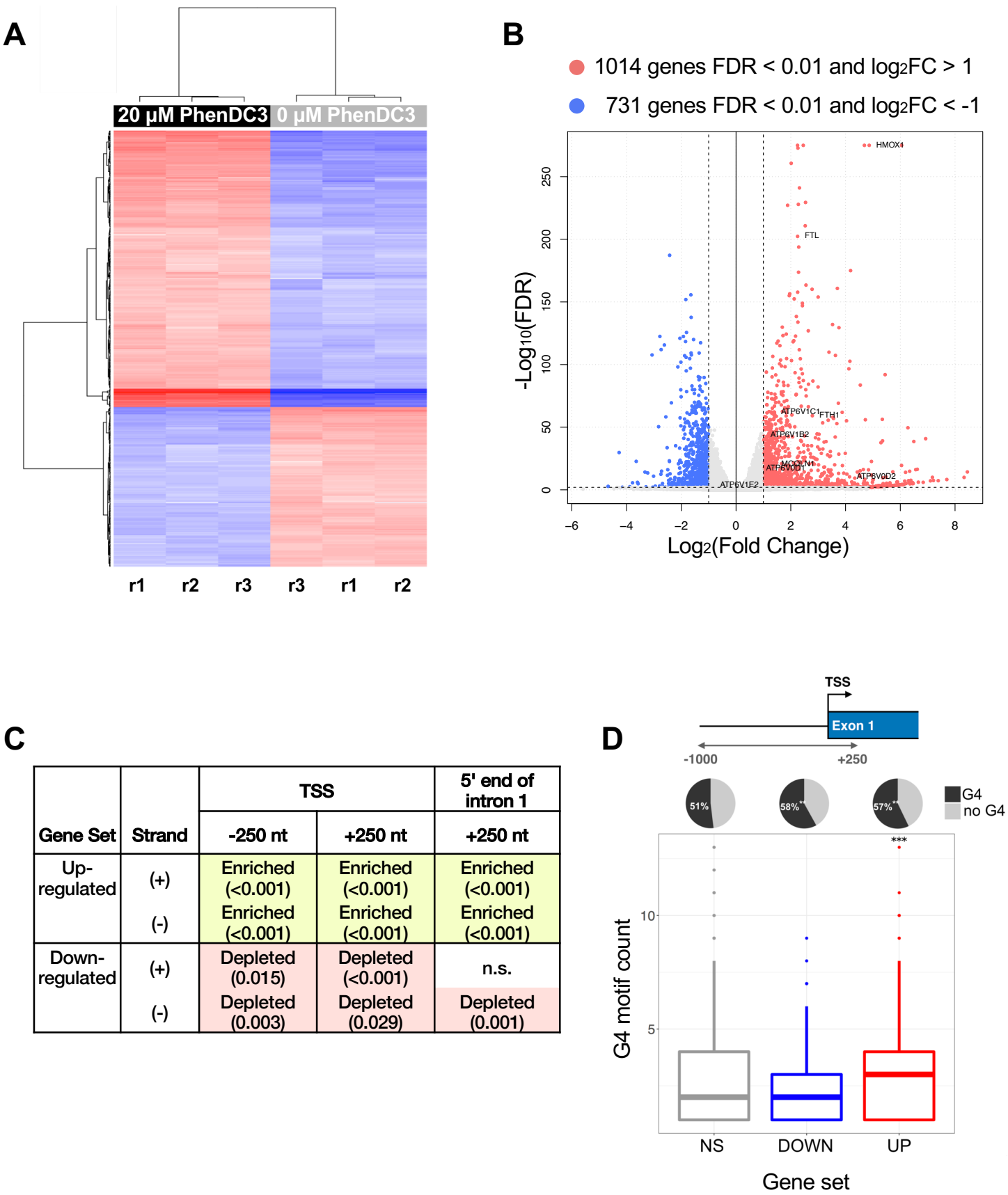

untreated
10 μM PhenDC3
20 μM PhenDC3
% cells
40
30
20
10
0
80.4
19.6
84.7
15.3
78.5
21.5
0 50K 100K 150K 200K 250K
DAPI-A

Figure 3

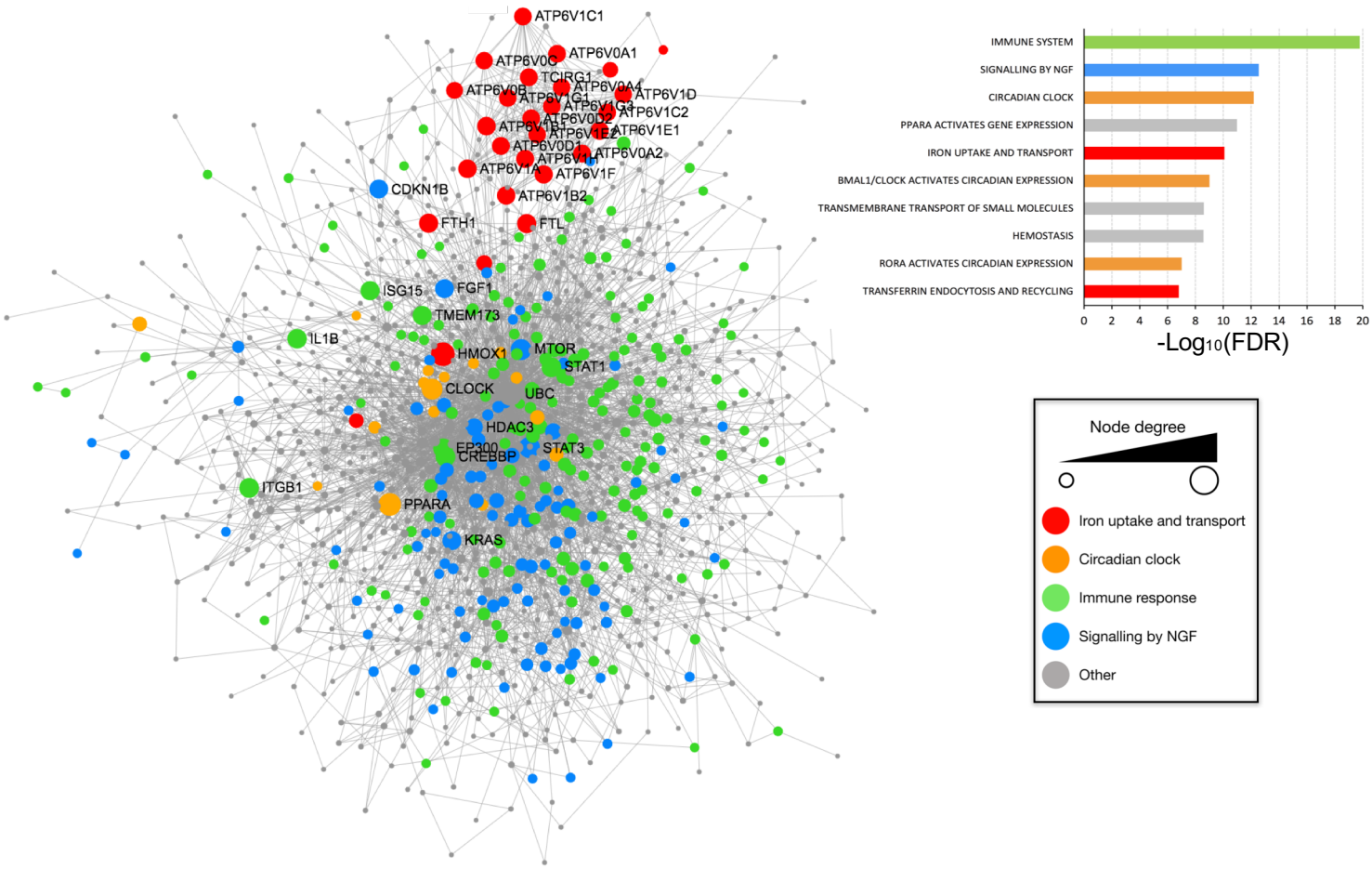


C

Gene Set	Strand	TSS -250 nt	TSS +250 nt	5' end of intron 1 +250 nt
Up-regulated	(+)	Enriched (<0.001)	Enriched (<0.001)	Enriched (<0.001)
	(-)	Enriched (<0.001)	Enriched (<0.001)	Enriched (<0.001)
Down-regulated	(+)	Depleted (0.015)	Depleted (<0.001)	n.s.
	(-)	Depleted (0.003)	Depleted (0.029)	Depleted (0.001)

Figure 4

A


ATP6V1C1
ATP6V0A1
ATP6V0C
TCIRG1
ATP6V0A4
ATP6V0B
ATP6V1D
ATP6V1G1
ATP6V1G3
ATP6V1C2
ATP6V0D2
ATP6V1E1
ATP6V1E2
ATP6V0D1
ATP6V0A2
ATP6V1H
ATP6V1A
ATP6V1F
CDKN1B
ATP6V1B2
FTH1
FTL
ISG15
FGF1
TMEM173
IL1B
HMOX1
MTOR
STAT1
CLOCK
UBC
HDAC3
EP300
CREBBP
STAT3
ITGB1
PPARA
KRAS
IMMUNE SYSTEM
SIGNALLING BY NGF
CIRCADIAN CLOCK
PPARA ACTIVATES GENE EXPRESSION
IRON UPTAKE AND TRANSPORT
BMAL1/CLOCK ACTIVATES CIRCADIAN EXPRESSION
TRANSMEMBRANE TRANSPORT OF SMALL MOLECULES
HEMOSTASIS
RORA ACTIVATES CIRCADIAN EXPRESSION
TRANSFERRIN ENDOCYTOSIS AND RECYCLING
0 2 4 6 8 10 12 14 16 18 20
-Log10(FDR)
Node degree
Iron uptake and transport
Circadian clock
Immune response
Signalling by NGF
Other

B

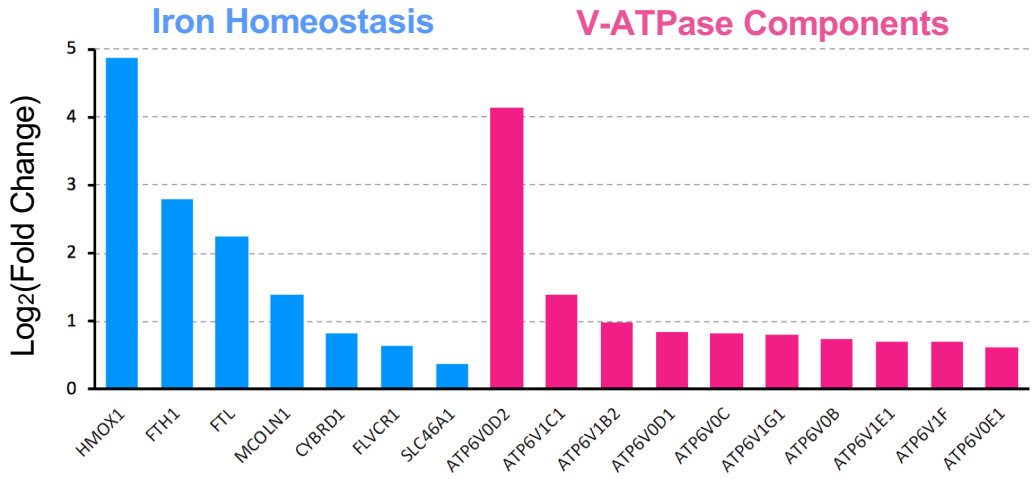
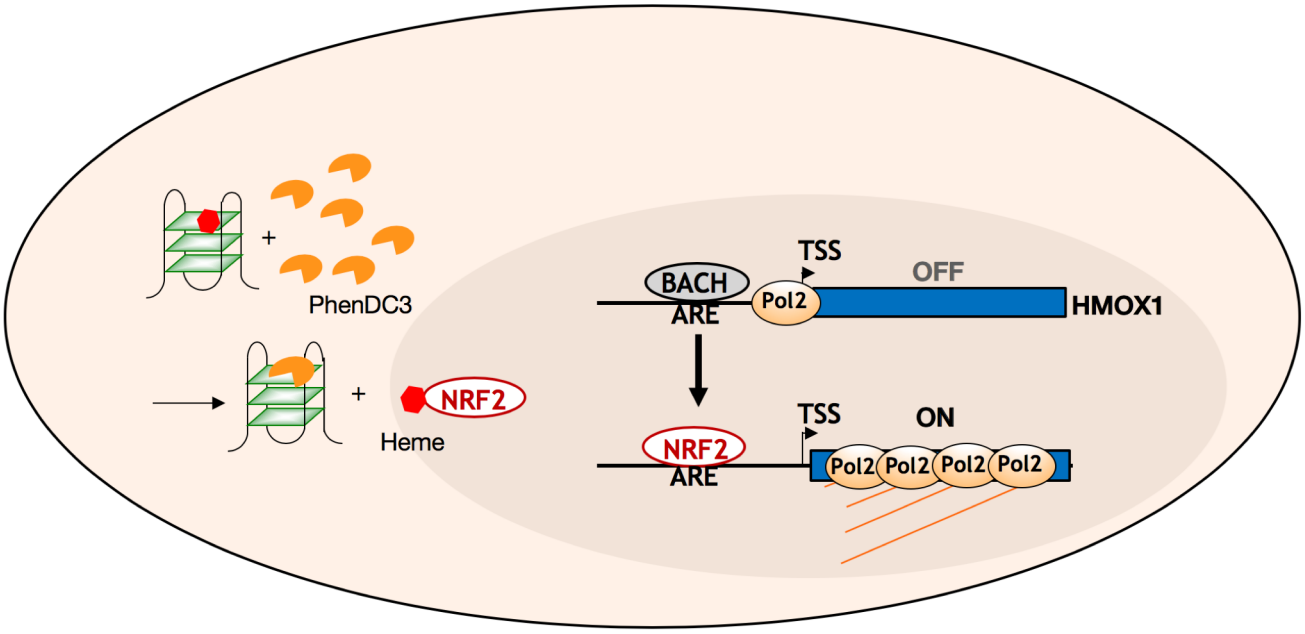

Iron Homeostasis
V-ATPase Components
Log2(Fold Change)
0
1
2
3
4
5
HMOX1
FTH1
FTL
MCOLN1
CYBRD1
FLVCR1
SLC46A1
ATP6V0D2
ATP6V1C1
ATP6V1B2
ATP6V0D1
ATP6V0C
ATP6V1G1
ATP6V0B
ATP6V1E1
ATP6V1F
ATP6V0E1

Figure 5


PhenDC3
Heme
NRF2
BACH
ARE
Pol2
TSS
OFF
HMOX1
NRF2
ARE
TSS
ON
Pol2
Pol2
Pol2
Pol2