HHS Public Access

Author manuscript

Nat Commun. Author manuscript; available in PMC 2015 June 10.

Published in final edited form as:

Nat Commun. ; 5: 5769. doi:10.1038/ncomms6769.

Regulation of tyrosine hydroxylase transcription by hnRNP K and DNA secondary structure

Kasturi Banerjee[1,3], **Meng Wang**[1,3], **Elizabeth Cai**[1], **Nana Fujiwara**[1], **Harriet Baker**[1,2], and **John W. Cave**[1,2]

[1]Burke Medical Research Institute, 785 Mamaroneck Ave, White Plains, NY, 10605

[2]Brain & Mind Research Institute, Weill Cornell Medical College, New York, NY, 10021

Abstract

Regulation of tyrosine hydroxylase gene (*Th*) transcription is critical for specifying and maintaining the dopaminergic neuronal phenotype. Here we define a molecular regulatory mechanism for *Th* transcription conserved in tetrapod vertebrates. We show that heterogeneous nuclear ribonucleoprotein (hnRNP) K is a transactivator of *Th* transcription. It binds to previously unreported and evolutionarily conserved G:C-rich regions in the *Th* proximal promoter. hnRNP K directly binds C-rich single DNA strands within these conserved regions and also associates with double-stranded sequences when proteins, such as CREB, are bound to an adjacent cis-regulatory element. The single DNA strands within the conserved G:C-rich regions adopt either G-quadruplex or i-motif secondary structures. We also show that small molecule-mediated stabilization of these secondary structures represses *Th* promoter activity. These data suggest that these secondary structures are targets for pharmacological modulation of the dopaminergic phenotype.

Dopamine is an essential neurotransmitter for numerous neurological functions that include locomotor behaviors, lactation, light/dark visual adaptation and odor detection. Consistent with its pleiotropic role in the nervous system, dopaminergic dysfunction in humans underlies many debilitating conditions, including Parkinson's Disease, schizophrenia, depression, drug addiction and attention deficit disorders. These important roles in neurobiology have motivated numerous studies on the molecular mechanisms responsible for development and maintenance of the dopaminergic phenotype. Many studies have concentrated on transcription regulatory mechanisms of *tyrosine hydroxylase* (*Th*), which encodes the rate-limiting enzyme in catecholamine neurotransmitter biosynthesis. Despite these intense efforts, the transcription factors and molecular mechanisms necessary to generate the complex patterns of *Th* transcription in the nervous system are only partially understood.

To whom correspondence should be addressed: John W. Cave, Burke Medical Research Institute, Weill Cornell Medical College, 785 Mamaroneck Ave, White Plains, NY, 10605, USA, Tel: 914-368-3140, Fax: 914-597-2821, joc2042@med.cornell.edu.

[3]Authors contributed equally to this work

Several studies have demonstrated that a cAMP response element (CRE) within the *Th* proximal promoter region is necessary for mediating transcription in catecholaminergic neurons.[1–3] In humans, single nucleotide polymorphisms within this site are associated with the neurometabolic disorder Tyrosine Hydroxylase Deficiency.[4] Although necessary, the CRE is not sufficient to establish the complex spatial expression pattern of *Th* in the brain.[5] NURR1 (NR4A2) is a region-specific transcription factor necessary for *Th* transcription in midbrain dopaminergic neurons.[6,7] The functionality of a putative NURR1 binding site[8,9] in the *Th* proximal promoter is difficult to confirm, however, since it partially overlaps with the TATA box. ETS-domain transcription factors have also been proposed to direct brain region-specific *Th* expression.[10] We previously showed that the ETS-domain transcription factor ER81 (ETV1) binds the *Th* promoter in mouse olfactory bulb (OB) dopaminergic neurons, but this molecular mechanism for regulating *Th* expression is rodent-specific.[11] Given the extensive use of rodents to model human dopaminergic systems, a high priority is to identify transcriptional regulatory mechanisms that are conserved between species.[12]

To identify functional and evolutionarily conserved transcriptional cis-regulatory elements, we align *Th* proximal promoter nucleotide sequences from a wide range of vertebrate species. This analysis identifies two novel, conserved G:C-rich regions upstream of the CRE that facilitate *Th* promoter activity. We show that these regions are bound by heterogeneous nuclear ribonucleoprotein K (hnRNP K) and adopt G-quadruplex and i-motif secondary structures. We also show that small molecule-mediated stabilization of these secondary structures represses *Th* promoter activity. Together, these findings reveal a novel regulatory mechanism for *Th* transcription conserved in most vertebrate species and suggest that secondary structures in the *Th* promoter are novel targets for pharmacological modulation of the dopaminergic phenotype.

Results

Conserved functional regions in the Th proximal promoter

To identify evolutionarily conserved regions, *Th* proximal promoter nucleotide sequences from species in five vertebrate orders (mammals, avians, reptiles, amphibians and fish) were aligned. Focusing on the region ~200 base pairs upstream of the transcription start site, the alignment revealed that all tetrapods (mammals, avians, reptiles and amphibians) contain a CRE and TATA box as well as two previously unrecognized G:C-rich regions upstream from the CRE (GC-R1 and GC-R2; Figure 1A). The length of these regions varied by species, but both contained several short motifs that were highly conserved.

The alignment of the *Th* proximal promoter in several fish species revealed similarities to tetrapods. Both a CRE and TATA box motif were identified in all of the fish species examined (Figure 1B). In contrast to tetrapods, however, a G:C-rich region was observed in only a subset of fish species. The G:C-rich region in fish resembled GC-R1 in tetrapods, including the partial conservation of a 5′-GGTGG-3′ sequence.

To determine whether the G:C-rich regions modulated *Th* promoter activity, luciferase transcription assays were performed using the 4.5kb upstream rat *Th* promoter in a *Th*-expressing cell line derived from the mouse OB. The rat 4.5kb upstream region was used

because it is the minimum sequence sufficient for directing *Th* promoter activity in all catecholaminergic brain regions in transgenic mice.[13] Mutations that overlapped highly conserved residues in both G:C-rich regions substantially reduced reporter gene expression in the OB cell line, but not as much as mutations to the CRE (Figure 1C and D). Mutations within the G:C-rich regions that did not overlap with highly conserved sequences (e.g. GC1mC) had little effect on promoter activity. Together, these findings indicate that highly conserved sequences in both G:C-rich regions are functional and important for activating *Th* transcription.

Identification of hnRNP K as a regulator of Th transcription

To identify proteins that bind to the G:C-rich regions, pull-down experiments were performed using biotinylated double-stranded oligonucleotides containing either the wild-type or mutated sequences. Initial studies focused on the 5′-GGTGG-3′ sequence in GC-R1 (Figure 2A) since this motif was strongly conserved in vertebrates (Figure 1A and 1B). The target oligonucleotides were incubated with nuclear lysate from *Th*-expressing cell lines derived from either the substantia nigra (SN) or the OB. Coomassie-stained gels showed a band that was differentially pulled down in both cell lines (Figure 2A). Mass spectrometry established that this band corresponded to hnRNP K. Transcription assays in human SH-SY5Y cells showed that luciferase activity under the control of the 4.5kb upstream rat *Th* promoter was up-regulated when hnRNP K was over-expressed (Figure 2B). Human *Th* promoter activity in SH-SY5Y cells was also up-regulated in SH-SY5Y cells (Supplementary Fig. 1).

Double-label immunofluorescence was used to establish *in vivo* co-expression of hnRNP K with TH in OB and SN dopaminergic neurons in both mice and humans. This analysis showed that nearly all TH-expressing cells co-expressed hnRNP K in both the mouse SN (217 out of 226 SN TH+ cells counted; Figure 2C–F) and the OB (131 out of 141 TH+ cells counted; Figure 2G–J). In tissue obtained from human subjects, most cells with TH also co-expressed hnRNP K in both the SN (159 out of 167 TH+ cells counted; Figure 2K–N) and OB (134 out of 145 TH+ counted; Figure 2O–R). In all mouse and human sections examined, there were hnRNP K-expressing cells lacking TH co-expression. Thus, the immunofluorescence data are consistent with a role for hnRNP K in regulating *Th* transcription, but the presence of hnRNP K+ cells lacking TH co-expression indicates that hnRNP K is not sufficient to induce *Th* transcription.

To test whether hnRNP K associated with the *Th* promoter in a cellular context, we performed chromatin immunoprecipitation (ChIP) assays with SH-SY5Y cells. These studies showed that hnRNP K directly bound the G:C-rich regions of the *Th* promoter (Figure 3A). Previous studies have shown that hnRNP K preferentially binds C-rich single-stranded DNA sequences.[14] To establish whether hnRNP K binds single DNA strands from the G:C-rich regions in the *Th* promoter, pull-down assays used single-stranded biotinylated oligonucleotides incubated with nuclear lysate from either OB or SN cell lines. The target oligonucleotide sequences contained either the wild-type G:C-rich region sequences or mutations overlapping highly conserved positions (Figure 3B). Western blots showed that

hnRNP K was preferentially pulled down by the wild-type C-rich strands of both GC-R1 and GC-R2 (Figure 3B; Supplementary Fig. 2).

On some gene promoters, hnRNP K binds double-stranded DNA sequences.[15,16] Electromobility shift assays (EMSAs) with a double-stranded oligonucleotide of the *Th* promoter, however, did not detect any binding of hnRNP K (lanes 1–4 in Figure 3C; Supplementary Fig. 3). The inability of hnRNP K to bind this duplex oligonucleotide was unexpected since this was the same duplex used initially to isolate hnRNP K in the protein pull-down experiment (Figure 2A). An important difference between these experiments, however, was that the protein pull-down experiments used nuclear lysate from *Th*-expressing cell lines, but the EMSA used purified hnRNP K protein. These data suggested that an additional factor in the nuclear lysate enabled hnRNP K to associate with the double-stranded DNA. Given the proximity of the CRE, we tested whether the presence of purified CRE Binding protein (CREB) influenced the binding behavior of hnRNP K. Additional electromobility shift assays showed that hnRNP K super-shifted CREB/DNA complexes (lanes 5–8 in Figure 3C; Supplementary Fig. 3). Both the CREB-only and CREB/hnRNP K samples produced multiple shifted bands, suggesting multiple protein/DNA complexes were present. CREB bound the probe even when the CRE was mutated (Supplementary Fig. 4), which suggested that the multiple protein/DNA complexes in Figure 3C were generated by CREB recognizing non-canonical, but as yet unidentified, binding sites in the *Th* promoter. The ability of hnRNP K to super-shift these complexes indicated that proteins bound to the CRE were required to recruit hnRNP K to the duplex *Th* proximal promoter. To confirm that the CRE was necessary for this recruitment, protein pull-down assays with OB and SN cell line nuclear lysate were performed using a duplex probe derived from GC-R1 with the CRE mutated (Figure 2A). These pull-down experiments differentially isolated a protein that mass spectrometry confirmed was hnRNP K, indicating that mutation of the CRE was sufficient to block pull-down of hnRNP K from nuclear lysates of both OB and SN cell lines (Figure 2A).

DNA secondary structure in the Th proximal promoter

Individual strands of gene promoter regions targeted by hnRNP K can adopt DNA secondary structures that modulate transcription.[17,18] C-rich strands can form i-motif structures and the opposing G-rich strands can generate G-quadruplexes.[19] To address whether the *Th* promoter can also adopt these DNA secondary structures, we examined single-stranded oligonucleotides from GC-R1 and GC-R2 by circular dichroism (CD) spectroscopy. Previous studies have established that i-motifs and G-quadruplexes are stabilized *in vitro* by low pH (<5) and high K+ concentrations (≥100mM), respectively.[20,21] Both wild-type G-rich strands from the *Th* promoter had similar CD absorption spectra with minima and a maxima that were enhanced in the presence of 100mM KCl (Figure 4A and 4B). These spectra also strongly resemble previously reported CD spectra for parallel-stranded G-quadruplexes.[22–26] In contrast to the wild-type sequences, mutations to the G-rich sequences from the *Th* promoter shifted the signal minima and maxima to longer wavelengths and blocked KCl-mediated enhancement of the signal maximum intensity (Figure 4A and 4B). CD spectra for both wild-type C-rich strands from the *Th* promoter were consistent with formation of i-motif secondary structures. Lowering the sample pH

from 7.6 to 4.6 shifted spectral minimum and maximum to longer wavelengths and strongly enhanced the intensity of the signal maximum (Figure 4C and 4D). These spectra were nearly identical to previous analyses of i-motif structures by CD spectroscopy.[23,24,27,28] Mutation of the C-rich regions in the *Th* promoter sequences blocked the pH-induced changes in the CD spectra (Figure 4C and 4D). Together, the CD spectra strongly indicate that the single-stranded sequences of the *Th* promoter adopt either G-quadruplex or i-motif secondary structures.

To establish whether DNA secondary structure was present on the *Th* promoter *in vivo*, ChIP experiments were performed with the BG4 antibody, which specifically recognizes G-quadruplex structures[29] (Supplementary Fig. 5). In both mouse OB and midbrain tissue, these studies revealed the presence of G-quadruplex structures on the *Th* promoter *in vivo* (Figure 5A and 5B). ChIP experiments also detected G-quadruplexes on the *Th* promoter in SH-SY5Y cells (Figure 5C).

The porphyrin compound TMPyP4 (Figure 6A) can stabilize both G-quadruplexes and i-motifs by non-intercalative mechanisms.[30,31] By contrast, the structural isomer, TMPyP2 (Figure 6A), is ineffective at stabilizing these DNA secondary structures.[32–34] To establish whether TMPyP4 can stabilize DNA secondary structures formed by the *Th* promoter DNA, oligonucleotide template strands derived from GC-R1 and GC-R2 were amplified by PCR in presence of either TMPyP4 or TMPyP2. The presence of TMPyP4, but not TMPyP2, was expected to diminish transcription by stabilizing the DNA secondary structure on the template strands and blocking the DNA polymerase enzyme. These assays showed that increasing TMPyP4 concentrations progressively reduced transcription of the wild-type template strands (Figure 6B). By contrast, the TMPyP2 only weakly reduced or even enhanced transcription product formation rates. Transcription of the mutated template strands also was not significantly impeded by TMPyP4 (Figure 6B). Together, these findings indicate that TMPyP4 can stabilize DNA secondary structures in the *Th* promoter.

Stabilization of DNA secondary structure on the *Th* promoter by TMPyP4 would be expected to preclude proteins, such as hnRNP K, from binding the single-stranded promoter DNA. Consistent with this mechanism, ChIP experiments conducted with SH-SY5Y cells showed reduced hnRNP K occupancy on the *Th* promoter when the cells were treated with TMPyP4 (Figure 6C).

To address whether small molecules, such as TMPyP4, can modulate *Th* transcription levels in a cellular context, we analyzed the effect of these compounds in depolarized forebrain organotypic slice cultures from a transgenic *Th*-GFP mouse strain expressing GFP under the control of the 9kb upstream region of the *Th* promoter.[35] Compared to either TMPyP2-treated or control sections, GFP expression was strongly attenuated in the main OB dopaminergic neurons of TMPyP4-treated sections (Figure 6D). These findings show that, when provided access to the brain, pharmacological agents that stabilize G-quadruplex and i-motif DNA secondary structures, such as TMPyP4, can modulate *Th* expression levels. Interestingly, TMPyP4 did not substantially disrupt GFP expression in the accessory OB (AOB; Figure 6D). The GFP-expressing cells in the AOB are unlikely catecholaminergic

neurons since they do not express TH protein (Supplementary Fig. 6), which suggests that *Th* is differentially regulated in non-catecholaminergic cells.

Discussion

This study identified a novel transcription regulatory mechanism on the *Th* promoter mediated by hnRNP K and DNA secondary structure. These studies showed that hnRNP K is a transactivator of *Th* transcription that directly binds C-rich single-stranded DNA sequences within conserved G:C-rich regions of the *Th* promoter. hnRNP K is an established transactivator targeting C-rich single-stranded DNA sequences of several oncogenes.[15,16,18,36] To our knowledge, however, this is the first study showing that hnRNP K also targets neurotransmitter biosynthesis genes. These studies also found that hnRNP K can associate with duplex DNA sequences from the *Th* promoter, but only when proteins, such as CREB, are bound to the CRE. Previous reports also demonstrated hnRNP K binding to DNA duplexes[15,16], but these studies used either whole or partially purified nuclear extracts and it is unclear whether additional proteins within the extracts assisted hnRNP K binding to the target duplex sequences. By contrast, the EMSAs reported here used only purified stocks of hnRNP K and CREB proteins. Thus, the DNA-binding synergy was not mediated by third party proteins bridging hnRNP K and CREB. These findings suggest an interaction between hnRNP K and CREB, but protein immunoprecipitation assays have failed to detect any evidence of a specific physical interaction between these two proteins (Supplementary Fig. 7). The inability to co-immunoprecipitate the proteins, however, may be because the interaction between these proteins masks the epitopes necessary for such experiments to detect them. Alternatively, deformations from canonical B-form duplex DNA structure induced by basic-leucine-zipper (bZIP) DNA binding domain proteins, such as CREB,[37,38] may be specifically recognized by hnRNP K and/or integral to the stability of hnRNP K/dsDNA complexes.

Together, hnRNP K and bZIP proteins likely co-regulate *Th* transcription in most dopaminergic neurons. The co-expression of hnRNP K with *Th* promoter in both the OB and SN suggests that its regulatory function is not region-specific. Previous studies have established that the CRE in the *Th* proximal promoter is required for *Th* transcription in all catecholaminergic neurons.[1–3,39,40] The specific bZIP protein complexes bound to the CRE, however, may be region-specific. We showed CREB bound the *Th* promoter in this study, but other bZIP proteins, including Fos-B and Jun-D, can also bind the CRE.[41] Both hnRNP K and bZIP proteins also contact the general transcription protein machinery. hnRNP K physically interacts with the TBP subunit of TFIID while bZIP proteins recruit co-activators, such as CBP/p300, that interact with several subunits of the RNA polymerase II protein complex.[42–44] Thus, both bZIP and hnRNP K proteins are important regulators of the dopaminergic neuronal phenotype through their interactions on the *Th* promoter.

This study also revealed that individual strands from the conserved G:C-rich regions in the *Th* promoter adopt either G-quadruplex or i-motif secondary structures. This mirrors previous studies showing that these types of DNA secondary structures form on G:C-rich regions of oncogene promoters also targeted by hnRNP K.[17,18] The conserved function of hnRNP K and DNA secondary structure on both *Th* and oncogene promoters leads us to

propose that the *Th* proximal promoter regulates transcription by a mechanism similar to what has been suggested for the *c-myc* oncogene[17] (Figure 7). On the *Th* promoter, the recruitment of bZIP proteins and their co-activators to the CRE facilitates assembly of the RNA Polymerase transcription machinery. Melting of core promoter duplex DNA by the RNA Polymerase complex generates negative supercoiling upstream of the complex that can promote separation of duplex DNA strands. The resulting single strands can adopt secondary structures, such as G-quadruplexes and i-motifs, which help dissipate the torsional stress.[45] These DNA secondary structures, especially G-quadruplexes, are associated with transcriptional repression through mechanisms that are not established, but may include either disruption of transcription elongation, blocking the binding of transcription activator proteins, or recruitment of transcription repressor proteins.[37,38] By contrast, binding of the single DNA strands by proteins such as hnRNP K prevents secondary structure formation and promotes gene transcription by contacting general transcription factors, such as TBP.[42,43] The loss of hnRNP K function *in vivo* would be expected to reduce *Th* transcription levels due both to the lost interaction with TBP and the increased rate of secondary structure formation. The lack of hnRNP K knock-out mouse strains, however, has precluded direct testing of this prediction. Alternative strategies to knock down hnRNP K by RNA interference in several neural cell lines using either transfection or lentiviral transduction have been impeded by hnRNP K's regulation of proliferation and survival genes.[46]

On the *c-myc* promoter, the CNBP and hnRNP K proteins stabilize the G-rich and C-rich single strands, respectively.[17] In both the OB and SN, we show that the C-rich sequences are also targeted by hnRNP K. Given the broad expression pattern of hnRNP K in the nervous system, we expect that the role of hnRNP K in regulating *Th* expression is general and not region-specific. The factors bound to the G-rich sequences on the *Th* promoter have not been established, but in a subset of dopaminergic neurons in both the OB and SN, the G-rich regions may be stabilized by hnRNP A1 (Supplementary Fig. 8–10). Further testing will establish whether secondary structures on the *Th* promoter are destabilized by proteins, such as NM23-H2, as they are on the *c-myc* promoter.[17,45]

The ability of TMPyP4 to repress *Th* expression suggests that DNA secondary structures on the *Th* promoter are novel pharmacological targets for small molecules. This method of repressing *Th* transcription could assist stem cell engineering protocols to produce non-catecholaminergic neuronal phenotypes. Current pharmacological efforts are concentrated on developing compounds that stabilize the G-quadruplexes in order to repress oncogene expression and telomerase activity[47], but development of agents that destabilize these structures on *Th* promoter may benefit therapeutic strategies to treat dopamine-related disorders, such as TH deficiency and Parkinson's Disease. Recent studies indicate G-quadruplexes are pervasive throughout the human genome and present in many gene promoter regions[48], which raises the possibility that pharmacological manipulation of DNA secondary structure may benefit additional neurological conditions.

The conservation of the G:C-rich regions, CRE and TATA box in most vertebrate species indicates that the proteins bound to these conserved regions and DNA secondary structure are together part of an evolutionarily ancient mechanism regulating *Th* transcription. An

important future direction is to establish how this mechanism interacts with other transcription factors previously identified as regulators of *Th* transcription. Understanding how NURR1 integrates with this mechanism is of particular interest. NURR1 is necessary for the development of midbrain dopaminergic neurons and activates the expression of several genes necessary for this phenotype, including *Th* and *c-Ret*.[6,7,49,50] Interestingly, the *c-Ret* promoter can also adopt G-quadruplex and i-motif secondary structures[23] and suggests that similar mechanisms identified here for *Th* may be also functional on other genes necessary for specifying the dopaminergic phenotype.

Methods

Genomic sequence alignment

All *Th* proximal promoter genomic sequences were obtained from Ensembl.org (http://ensembl.org). For all species, 500bp sequences upstream from the translation start site were simultaneously aligned and visualized using Multi-LAGAN and mVISTA web-based services (http://genome.lbl.gov/vista/mvista/submit.shtml).[51,52] The Ensembl.org gene identifiers for *Th* in each species: rat, *Rattus norvegicus* ENSRNOG00000020410; mouse, *Mus musculus* ENSMUSG00000000214; guinea pig, *Cavia porcellus* ENSCPOG00000012591; human, *Homo sapiens* ENSG00000180176; Gorilla, *Gorilla gorilla* ENSGGOG00000003978; Orangutan, *Pongo abelii* ENSPPYG00000002924; cat, *Felis catus* ENSFCAG00000006939; dog, *Canis familiaris* ENSCAFG00000010099; cow, *Bos taurus* ENSBTAG00000026768; dolphin, *Tursiops truncatus* ENSTTRG00000015465; chicken, *Gallus gallus* ENSGALG00000006551; turkey, *Meleagris gallpavo* ENSMGAG00000007908; zebrafinch, *Taeniopygia guttata* ENSTGUG00000009295; anole, *Anois carolinensis* ENSACAG00000009753; frog, *Xenopus tropicalis* ENSXETG00000014030; puffer, *Takifugu rubripes* ENSTRUG00000009267; Stickleback, *Gasterosteus aculeatus* ENSGACG00000011104; Zebrafish, *Danio rerio* ENSDARG00000030621; Medaka, *Oryzias latipes* ENSORLG00000018964; Cod; *Gadus morhua* ENSGMOG00000017881; Coelancath, *Latimeria chalumnae* ENSLACG00000006317.

Cell culture

Generation of the SN-derived *Th*-expressing cell line (SN4741) has been previously described.[53] Briefly, this clonal cell line was established from mesencephalic tissue dissected from embryonic day 13.5 transgenic mice expressing the SV40 large T-antigen under the control of the rat 9kb *Th* upstream region. The OB *Th*-expressing cell line (OB4671) is also a clonal line derived from this transgenic mouse strain, but it was generated using OB tissue dissected from post-natal day 2 animals. Both OB and SN cell lines were maintained at 33°C with 95% air and 5% CO_2 on primaria-coated culture dishes in DMEM cell culture medium supplemented with 10% FBS and 1% penicillin/streptomycin. The human SH-SY5Y neuroblastoma cell line was maintained at 37°C with 95% air and 5% CO_2 in DMEM/F12 media supplemented with plus GlutaMax^{tm} (Life Technologies).

Luciferase transcription assays

Transcription assays in the OB cell line were conducted with stably transfected cells. For each line, 2.5×10^6 cells in 60mm Primaria-coated culture dishes were transfected using Fugene HD transfection reagent (Roche) with 2μg of pGL4.20 firefly luciferase reporter plasmid (Promega). The firefly luciferase was under the control of the rat 4.5kb *Th* upstream region containing either the wild-type or mutated G:C-rich regions. The pGL4.20 plasmid also constitutively expressed a puromycin resistance gene to enable selection of transfected cells. 48 hours after transfection, culture media was supplemented with 25μg/ml puromycin to select transfected cells. The puromycin concentration was reduced to 8.3μg/ml after one subculture and this reduced level was used to maintain the lines.

To measure luciferase activity in stably transfected cell lines, cultures were dissociated with trypsin, washed and resuspended in 1ml of PBS. The density (OD) of the isolated cells was measured at a wavelength of 600nm. Cells were then pelleted and treated with reporter lysis buffer (Promega). Using a firefly luciferase assay kit (Promega), luminescence was measured with LMaxII luminometer (Molecular Devices). Luminescence readings were normalized to the respective cell densities. Luciferase activity was reported as the mean of three independent transfection experiments with error bars representing the standard error of the mean. Significant changes in activity were assessed either by two-tailed Student's *t*-test or by ANOVA with appropriate post-hoc tests.

Transcription assays with SH-SY5Y cells over-expressing hnRNP K were conducted using transient transfections with Lipofectamine LTX transfection reagent (Life Technologies). These experiments used the pGL4.20 plasmid (Promega) with a firefly luciferase reporter gene under the control of the wild-type rat 4.5kb upstream region and a pGFP-hnRNP K expression plasmid (a generous gift from Dr. Rama Natarajan, Beckman Research Institute). To control for variations in transfection efficiency, cells were also transfected with pRL-CMV (Promega), which constitutively expressed Renilla luciferase. Cells were transfected according to the manufacturer's protocol and maintained for 48 hours before both firefly and Renilla luciferase activities were measured using the Dual-Glo Luciferase Assay System (Promega). Luminescence was measured with LMaxII luminometer (Molecular Devices). Luciferase activity was reported as the mean of three independent transfection experiments with error bars representing the standard error of the mean. Significant changes in activity were assessed either by two-tailed Students *t*-test or by ANOVA with appropriate post-hoc tests.

Protein pull-down and mass spectrometry analysis

Protein pull-down assays using biotinylated DNA oligonucleotides were based on a previous study with minor modifications.[54] *Th*-expressing OB and SN cell lines were grown to a density of 10^5cells/ml on 100mm primaria-coated culture dishes before being harvested and lysed in RIPA buffer supplemented with protease inhibitors. Cellular debris was cleared from the lysate using centrifugation and protein content was estimated by a Bradford Assay (Bio-Rad). Immobilized DNA oligonucleotides were prepared by incubating 20μg of annealed duplex oligonucleotides with 100 μg of streptavidin-coated Dynabeads (Life Technologies) in 400μl of buffer (10mM Tris, pH 7.5, 1mM EDTA, 1M NaCl, 0.003%

NP40) for 1 hour at room temperature with constant rotation. To minimize non-specific interactions, the immobilized oligonucleotides were incubated for 30 min at room temperature with 2.5mg/ml BSA in 10mM HEPES, pH 7.6, 100mM potassium glutamate, 2.5mM DTT, 10mM magnesium acetate, 5mM EGTA, 3.5% glycerol with 0.003% NP40 and 5 mg/ml polyvinylpyrrolidone. Cell lysate containing 100μg of protein was then incubated with immobilized oligonucleotides for 4 hours at 4°C with constant rotation in 400μl of 10mM HEPES, pH 7.6, 100mM potassium glutamate, 80mM KCl, 2.5mM DTT, 10mM magnesium acetate, 5mM EGTA, 3.5% glycerol with 0.001% NP40 and 1μg of non-specific carrier DNA. After four hours, the immobilized protein/DNA oligo complexes were washed three times with 300μl of 10mM HEPES, pH 7.6, 100mM potassium glutamate, 2.5mM DTT, 10mM magnesium acetate, 5mM EGTA, 3.5% glycerol, 0.5mg/ml BSA and 0.05% NP40. Proteins bound to the immobilized DNA were eluted in 50μl of SDS-PAGE sample buffer and resolved by electrophoresis on a NuPAGE 4–12% gradient Bis–Tris gel (Life Technologies). The gel was stained with Coomassie Brilliant Blue R-250 (Boston Bioproducts) overnight and destained with a solution of 30% methanol, 7% acetic acid and 63% water. Gels were digitally scanned for image archiving and bands of interest were identified for further analysis. Gels were sent to the Cornell University Life Sciences core mass spectrometry and proteomics facility. This core facility excised the bands of interest, performed in-gel proteolytic digestion and protein fragment extraction. Protein fragments were analyzed by MALDI-TOF mass spectrometry and compared to proteomic databases using Mascot 2.2 software to identify the parent protein.

The duplex oligonucleotides used for protein pull-down assays corresponded to rat *Th* genomic positions −115 to −64 (relative to the start of translation).

The wild-type forward strand (biotinylated): 5′-
GCAGGAGGTAGGAGGTGGGGGACCCAGAGGGGCTTTGACGTCAGCCTGGCC-3′

The wild-type reverse strand: 5′-
GGCCAGGCTGACGTCAAAGCCCCTCTGGGTCCCCCACCTCCTACCTCCTGC -3′

The GC-R1 mutant forward strand (biotinylated; mutated positions underlined): 5′-
GCAGGAGGTAGGCTACATCAGACCCAGAGGGGCTTTGACGTCAGCCTGGCC-3′

The GC-R1 mutant reverse strand (mutated positions underlined): 5′-
GGCCAGGCTGACGTCAAAGCCCCTCTGGGTCTGATGTAGCCTACCTCCTGC-3′

The CRE mutant forward strand (biotinylated; mutated positions underlined): 5′-
GCAGGAGGTAGGAGGTGGGGGACCCAGAGGGGCTTGTCATCAGGCCTGGCC-3′

The CRE mutant reverse strand (mutated positions underlined): 5′-
GGCCAGGCCTGATGACAAGCCCCTCTGGGTCCCCCACCTCCTACCTCCTGC-3′

hnRNP K pull-down assays

1μg of single-stranded biotinylated DNA oligonucleotide was incubated with 5ug of streptavidin-coated Dynabeads (Life Technologies) in 300μL of streptavidin-biotin binding buffer (10mM Tris, pH 7.5, 1mM EDTA, 1M NaCl, 0.003% NP40) at room temperature for

20 min. To minimize non-specific interactions, the oligo-bead complexes were incubated for 30 min with a blocking buffer (2.5mg/ml BSA in 10mM HEPES, pH 7.6, 100mM potassium glutamate, 2.5mM DTT, 10mM magnesium acetate, 5mM EGTA, 3.5% glycerol with 0.003% NP40 and 5 mg/ml polyvinylpyrrolidone). *Th*-expressing OB and SN cell lines were grown in 10mm dishes and harvested at 80% confluence. A total of 3×10^7 cells were lysed and the nuclei were extracted with NE-PER nuclear and cytoplasmic extraction reagents (Thermo Scientific). Immobilized oligonucleotides were incubated with 100μg of nuclear extract for one hour at 4°C with constant rotation in 400μl of protein binding buffer (10mM HEPES, pH 7.6, 100mM potassium glutamate, 80mM KCl, 2.5mM DTT, 10mM magnesium acetate, 5mM EGTA, 3.5% glycerol with 0.001% NP40 and 1μg of non-specific carrier DNA). Protein-DNA complexes were then washed three times (10mM HEPES, pH 7.6, 100mM potassium glutamate, 2.5mM DTT, 10mM magnesium acetate, 5mM EGTA, 3.5% glycerol, 0.5mg/ml BSA and 0.05% NP40). Proteins bound to the oligonucelotides were then eluted with 10μL of denaturing Laemmli sample loading buffer at 37°C for 15 min. hnRNP K protein in the supernatant was visualized by Western blot with a rabbit anti-hnRNP K antibody (Abcam) at 1:2000 dilution using an Odyssey imaging system (Li-Cor Biosciences). Band intensities were quantified using ImageJ software (National Institutes of Health) and are reported as the mean of duplicate pull-down assays with error bars representing the difference between the mean and the individual trials. The oligonucleotides sequences used in the experiments were:

wild-type GC-R1 forward strand, 5′-
GCAGGCGTGGAGAGGATGCGCAGGAGGTAGGAGGTGGGGGACCCAGAGGGGC
TTTGACG -3′

mutant GC-R1 forward strand (mutated positions underlined), 5′-
GCAGGCGTGGAGAGGATGCGCCATCATTAGGAATCATATGACCCAGAGGGGCT
TTGACG-3′

wild-type GC-R1 reverse strand, 5′-
CGTCAAAGCCCCTCTGGGTCCCCCACCTCCTACCTCCTGCGCATCCTCTCCACGC
CTGC-3′

mutant GC-R1 reverse strand (mutated positions underlined), 5′-
CGTCAAAGCCCCTCTGGGTCATATGATTCCTAATGATGGCGCATCCTCTCCACGC
CTGC-3′

wild-type GC-R2 forward strand, 5′-
TCTCGTCGCCCTCGCTCCATGCCCACCCCCGCCTCCCTCAGGCAC-3′

mutant GC-R2 forward strand (mutated positions underlined), 5′-
TCTCGTCGCCCTCGCTCCATGCATGATATATATGATATCAGGCAC -3′

wild-type GC-R2 reverse strand, 5′-
GTGCCTGAGGGAGGCGGGGGTGGGCATGGAGCGAGGGCGACGAGA-3′

mutant GC-R2 reverse strand (mutated positions underlined), 5′-GTGCCTGATATCATATATATCATGCATGGAGCGAGGGCGACGAGA -3′

Electrophoretic Mobility Shift Assay

Both duplex and single-stranded oligonucleotides were end-labeled with γ-P^{32} dATP using T4 polynucleotide kinase (New England Biolabs) and unincorporated nucleotides were removed using Bio-30 spin columns (Bio-Rad). All protein/DNA binding reactions were conducted in 20μL with 5mM HEPES (pH 7.9), 25mM KCl, 0.05mM EDTA, 0.125mM phenylmethysulfonyl fluoride, 1mg/mL of non-specific herring sperm carrier DNA, 0.0l% NP40 and 2μL of radiolabeled DNA probe. Binding reactions either used purified protein stocks of hnRNP K (Abcam) or CREB (generous gift from Dr. Jennifer K. Nyborg, Colorado State University). Binding reactions were carried out at 4°C for one hour before being resolved on either 3% or 5% 29:1 acrylamide/bis-acrylamide gels with 1/4x TBE buffer at 215V for two hours. Gels were dried and then imaged using a BAS2500 phosphoimager (GE Healthcare). The duplex oligonucleotides used for the EMSAs were the same as the wild-type sequences used for the protein pull-down assays and mass spectrometry analysis (see above).

Immunohistology and antibodies

Immunohistological analyses in mice used tissue from adult C57BL6 mice that were anesthetized with an overdose of pentobarbital (100mg/kg) and perfused transcardially with phosphate-buffered 4% formaldehyde. Brains were removed and post-fixed for four hours before being cryo-protected in 30% sucrose and cut in 30μm sagittal sections.

Human OB tissue was provided by the Harvard Brain Tissue Resource Center (HBTRC) and the human SN tissue was provided by the New York Brain Bank (NYBB)/Taub Institute for Research on Alzheimer's Disease and the Aging Brain. Tissue was provided from individuals without evidence of neurodegeneration, but the donor selection process was conducted exclusively by the HTBRC and NYBB. As such, all consent forms and patient information are maintained under protocols established by the HTBRC and NYBB, which are sanctioned by Institutional Review Boards approved by the National Institutes of Health. All human tissue was obtained post-mortem and formalin-fixed before being cryo-protected in 30% sucrose and cut in 30μm sagittal sections. Antigen retrieval was performed on all human tissue sections by immersing sections in 100mM Tris (pH 9) with 10mM glycine overnight at 75°C, then allowing sections to cool to room temperature before being washed with PBS.

For both mouse and human tissue, sections were blocked with 1% BSA in PBS for one hour before overnight incubation with primary antibody. The rabbit anti-hnRNP K (Abcam) primary antibody was used at a 1:500 dilution. The mouse TH primary antibody (Immunostar) was used at 1:1000. All primary antibodies were visualized with Alexa-Fluor secondary antibodies (Life Technologies) at a 1:400 dilution. Sections were imaged on either a Nikon 80i Eclipse fluorescence microscope or a Zeiss Meta 500 scanning laser confocal microscope.

Chromatin Immunprecipitation assays

Chromatin immunoprecipitation (ChIP) experiments to detect hnRNP K occupancy on the *Th* promoter used SH-SY5Y cells grown in 10mm culture dishes to 90% confluence before being washed with PBS, and then cross-linked with 1% formaldehyde in PBS at room temperature for nine minutes. For ChIP experiments probing the effect of either TMPyP4 or TMPyP2 exposure on hnRNP K occupancy, cells were treated with 100mM TMPyP4 or TMPyP2 for 24 hours prior to being washed and cross-linked with formaldehyde. Fixation was terminated by addition of 125mM glycine and the cells were washed twice with PBS before being collected with a sterile cell scraper. Cells were pelleted at 800xg centrifugation at 4°C for 5 min. The pellet was washed with PBS before lysing the cells with SDS Lysis Buffer (Millipore) and the chromatin was sheared with a Bioruptor (Diagenode). Immunoprecipitation of cross-linked protein/DNA complexes were performed using the Magna ChIP Protein-A/G kit (Millipore) following the manufacurer's instructions. Immunoprecipitation reactions used 4μg of either rabbit anti-hnRNP K antibody (Abcam) or normal rabbit IgG (Santa Cruz) and reverse cross-linking was done overnight in presence of Proteinase K at 62°C with shaking. Immunoprecipitated *Th* promoter genomic DNA fragments were amplified and quantitated using a 7500 Fast Real-time PCR System (Applied Biosystems) and SYBR Green PCR master mix (Applied Biosystems). Primer sequences used for amplifying the G:C-rich regions of the *Th* proximal promoter were 5′-CTGCGCCTCCCTCCTTCCTCAC-3′ and 5′-GAGCTGACGTCAAAGCCCCCTCTG-3′. For all hnRNP K ChIP experiments, three independent experiments were conducted and the mean relative enrichment of the *Th* proximal promoter is reported with error bars representing the standard error of the mean. Statistical significance was assessed using the two-tailed Student's t-test.

ChIP experiments to detect G-quadruplex secondary structures on the *Th* promoter used a His/FLAG-tagged BG4 single chain antibody.[29] An expression plasmid for this antibody was generously provided by Dr. Shankar Balasubramnian (Cambridge, UK). Control experiments used His-tagged thioredoxin expressed from a pET-32a expression plasmid (Novagen/EMD Millipore). Both His-tagged BG4 and Thioredoxin proteins were generated with BL21(DE3) *E. coli* cells. An overnight growth was used to inoculate 300mL of LB media to an OD_{600} of 0.1, which was then cultured at 37°C with shaking. When the OD_{600} reached 0.6, protein expression was induced with 1mM IPTG (Calbiochem). Six hours after induction, cells were spun down and the pellet was washed with PBS before being resuspended 1.6mL of equilibration buffer (PBS, with 10mM imidazole, pH 7.4, and 1mM PMSF). The resuspended cells were placed on ice, lysed by sonication and cellular debris was removed by centrifugation. Lysate with equivalent loads of either BG4 or Thioredoxin was incubated on HisPur Ni-NTA resin (Thermo Scientific) at 4°C for two hours, after which sheared chromatin from either SH-SY5Y cells, adult male and female *Th*-GFP mouse tissue from either the OB or midbrain tissue was added and incubated at 4°C overnight. The BG4/Thioredoxin-bound resin was washed 3 times with 1X PBS containing 25mM imidazole, and eluted in PBS with 1mM EDTA and 25mM imidazole. Samples were reverse cross-linked overnight at 62 °C on a shaker with 1μg of Proteinase K. DNA was purified using MinElute PCR purification columns (Qiagen) and *Th* promoter genomic DNA fragments were amplified and quantitated using a 7500 Fast Real-time PCR System

(Applied Biosystems) and SYBR Green PCR master mix (Applied Biosystems). Data are presented as the mean from three independent ChIP experiments with error bars representing the standard error of the mean. Statistical significance was assessed using two-tailed Student's t-tests. Primer sequences used to amplify the G:C-rich regions of the human *Th* proximal promoter (from SH-SY5Y cells) were: 5′-CTGCGCCTCCCTCCTTCCTCAC-3′ and 5′-GAGCTGACGTCAAAGCCCCCTCTG-3′. Primer sequences used to amplify the G:C-rich regions of the mouse *Th* proximal promoter (from OB and midbrain tissue) were: 5′-GTCGCCCTCGCTCTGTGCCCA-3′ and 5′-GGCTGACGTCAAAGCCCCTCTGG-3′.

Circular dichroism experiments

Circular dichroism (CD) experiments were performed with an Aviv 414 CD spectrophotomer. All spectra were collected with a wavelength range from 200–330 nm, 1nm step widths, and 1 second dwell times. Displayed spectra for each sample are an average of three scans. DNA oligonucleotide sequences used for these studies were identical to those used for the hnRNP K pull-down assays (see above) and all were used at concentrations of 5uM. For both wild-type and mutant G-rich sequences from GC-R1 and GC-R2, samples were in Tris buffer at pH 7.4 with either 0 or 100mM KCl. For both wild-type and mutant C-rich sequences from GC-R1 and GC-R2, samples were in sodium citrate/phosphate buffer at either pH 7.6 or 4.6.

PCR interference assays with TMPyP4 and TMPyP2

All PCR reactions were performed on a 7500 Fast Real-time PCR System (Applied Biosystems) using SYBR Green PCR master mix (Applied Biosystems) with 0, 5, 10 or 25nM of either TMPyP4 (Sigma-Aldrich) or TMPyP2 (Frontier Scientific). The data shown are the mean of three independent trials with error bars representing the standard error of the mean. Both wild-type and mutant oligonucleotide template strands were the same as those used in the hnRNP K pull-down experiments (see above).

Primer annealing to the GC-R1 wild-type and mutant forward strand: 5′-AGTCAGTCCGTCAAAG-3′

Primer annealing to the GC-R1 wild-type and mutant reverse strand: 5′-AGTCAGTCGCAGGCGT-3′

Primer annealing to the GC-R2 wild-type and mutant forward strand: 5′-AGTCAGTCGTGCCTGA-3′

Primer annealing to the GC-R2 wild-type and mutant reverse strand: 5′-AGTCAGTCTCTCGTCG-3′

Animals

Transgenic *Th*-GFP mice were obtained from Dr. Kazuto Kobayashi.[35] All mice were housed in humidity-controlled cages at 22 °C under a 12:12 hour light/dark cycle and provided with food and water *ad libitum*. All procedures were carried out under protocols approved by the Weill Cornell Medical College Institutional Animal Care and Use Committee and conformed to NIH guidelines.

Organotypic slice cultures

Whole brains from male and female *Th*-GFP mouse pups ranging in age from postnatal day 3 or 4 were dissected and placed into Petri dishes containing ice-cold dissection medium (Leibovitz's L-15 media, pH 7.4; Life Technologies). Brains were transected mid-sagittally and each hemisphere was placed with the medial face down into a plastic tissue embedding mold (Polysciences) and embedded in warm (37 °C) 4% low- melting-point agarose (Sigma) in L-15 medium, then rapidly cooled on ice and cut into 200 μm sagittal slices with a vibratome (Model 1000; Ted Pella Inc.). Slices were cultured in depolarizing conditions (25mM KCl) for 24 hours at 37°C and 5% CO_2 in Neurobasal medium (Invitrogen) supplemented with B27 (1:50, Invitrogen), 0.5 mM L-glutamine (Invitrogen) and 25 μg/ml gentamycin (Sigma). The slices were also treated either with100μM TMPyP4, 100μM TMPyP2 or no drug. All slices were imaged with a Nikon Eclipse 80i fluorescence microscope at 0 and 24 hours.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Rama Natarajan (Beckman Research Institute) for providing the pGFP-hnRNP K expression plasmid, Dr. Jennifer K. Nyborg (Colorado State University) for providing purified CREB protein, Dr. David Eliezer (Weill Cornell Medical College) for his assistance with the CD spectroscopy, Dr. Shakar Balasubramanian (Cambridge University) for the BG4 antibody, and both the Harvard Brain Tissue Resource Center and New York Brain Bank at Columbia University for human brain tissue. This work was supported by NIH grant R01 DC008955 and the Burke Medical Research Institute.

References

1. Kim KS, Lee MK, Carroll J, Joh TH. Both the basal and inducible transcription of the tyrosine hydroxylase gene are dependent upon a cAMP response element. J Biol Chem. 1993; 268:15689–95. [PubMed: 8101843]
2. Lewis-Tuffin LJ, Quinn PG, Chikaraishi DM. Tyrosine hydroxylase transcription depends primarily on cAMP response element activity, regardless of the type of inducing stimulus. Mol Cell Neurosci. 2004; 25:536–47. [PubMed: 15033181]
3. Trocme C, et al. CRE and TRE sequences of the rat tyrosine hydroxylase promoter are required for TH basal expression in adult mice but not in the embryo. Eur J Neurosci. 1998; 10:508–21. [PubMed: 9749713]
4. Willemsen MA, et al. Tyrosine hydroxylase deficiency: a treatable disorder of brain catecholamine biosynthesis. Brain. 2010; 133:1810–22. [PubMed: 20430833]
5. Min N, Joh TH, Kim KS, Peng C, Son JH. 5′ upstream DNA sequence of the rat tyrosine hydroxylase gene directs high-level and tissue-specific expression to catecholaminergic neurons in the central nervous system of transgenic mice. Brain Res Mol Brain Res. 1994; 27:281–9. [PubMed: 7898312]
6. Saucedo-Cardenas O, et al. Nurr1 is essential for the induction of the dopaminergic phenotype and the survival of ventral mesencephalic late dopaminergic precursor neurons. Proc Natl Acad Sci U S A. 1998; 95:4013–8. [PubMed: 9520484]
7. Zetterstrom RH, et al. Dopamine neuron agenesis in Nurr1-deficient mice. Science. 1997; 276:248–50. [PubMed: 9092472]
8. Iwawaki T, Kohno K, Kobayashi K. Identification of a potential nurr1 response element that activates the tyrosine hydroxylase gene promoter in cultured cells. Biochem Biophys Res Commun. 2000; 274:590–5. [PubMed: 10924322]

9. Kim KS, et al. Orphan nuclear receptor Nurr1 directly transactivates the promoter activity of the tyrosine hydroxylase gene in a cell-specific manner. J Neurochem. 2003; 85:622–34. [PubMed: 12694388]
10. Flames N, Hobert O. Gene regulatory logic of dopamine neuron differentiation. Nature. 2009; 458:885–9. [PubMed: 19287374]
11. Cave JW, et al. Differential regulation of dopaminergic gene expression by Er81. J Neurosci. 2010; 30:4717–24. [PubMed: 20357122]
12. Cave, JW.; Banerjee, K.; Baker, H. Species-specific molecular mechanisms establishing the dopamine neuronal phenotype. In: Kudo, E.; Fuji, Y., editors. Dopamine: Functions, Regulation and Health Effects. Nova Science Publishers, Inc; Happauge, NY: 2012. p. 195-214.
13. Schimmel JJ, Crews L, Roffler-Tarlov S, Chikaraishi DM. 4.5 kb of the rat tyrosine hydroxylase 5′ flanking sequence directs tissue specific expression during development and contains consensus sites for multiple transcription factors. Brain Res Mol Brain Res. 1999; 74:1–14. [PubMed: 10640671]
14. Tomonaga T, Levens D. Heterogeneous nuclear ribonucleoprotein K is a DNA-binding transactivator. J Biol Chem. 1995; 270:4875–81. [PubMed: 7876260]
15. Ritchie SA, et al. Identification of the SRC pyrimidine-binding protein (SPy) as hnRNP K: implications in the regulation of SRC1A transcription. Nucleic Acids Res. 2003; 31:1502–13. [PubMed: 12595559]
16. Takimoto M, et al. Specific binding of heterogeneous ribonucleoprotein particle protein K to the human c-myc promoter, in vitro. J Biol Chem. 1993; 268:18249–58. [PubMed: 8349701]
17. Brooks TA, Kendrick S, Hurley L. Making sense of G-quadruplex and i-motif functions in oncogene promoters. FEBS J. 2010; 277:3459–69. [PubMed: 20670278]
18. Uribe DJ, Guo K, Shin YJ, Sun D. Heterogeneous nuclear ribonucleoprotein K and nucleolin as transcriptional activators of the vascular endothelial growth factor promoter through interaction with secondary DNA structures. Biochemistry. 2011; 50:3796–806. [PubMed: 21466159]
19. Sun D, Hurley LH. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. J Med Chem. 2009; 52:2863–74. [PubMed: 19385599]
20. Gehring K, Leroy JL, Gueron M. A tetrameric DNA structure with protonated cytosine. cytosine base pairs. Nature. 1993; 363:561–5. [PubMed: 8389423]
21. Burge S, Parkinson GN, Hazel P, Todd AK, Neidle S. Quadruplex DNA: sequence, topology and structure. Nucleic Acids Res. 2006; 34:5402–15. [PubMed: 17012276]
22. Cogoi S, Xodo LE. G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. Nucleic Acids Res. 2006; 34:2536–49. [PubMed: 16687659]
23. Guo K, Gokhale V, Hurley LH, Sun D. Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. Nucleic Acids Res. 2008; 36:4598–608. [PubMed: 18614607]
24. Guo K, et al. Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene. Journal of the American Chemical Society. 2007; 129:10220–8. [PubMed: 17672459]
25. Mathad RI, Hatzakis E, Dai J, Yang D. c-MYC promoter G-quadruplex formed at the 5′-end of NHE III1 element: insights into biological relevance and parallel-stranded G-quadruplex stability. Nucleic Acids Res. 2011; 39:9023–33. [PubMed: 21795379]
26. Qin Y, Rezler EM, Gokhale V, Sun D, Hurley LH. Characterization of the G-quadruplexes in the duplex nuclease hypersensitive element of the PDGF-A promoter and modulation of PDGF-A promoter activity by TMPyP4. Nucleic Acids Res. 2007; 35:7698–713. [PubMed: 17984069]
27. Dettler JM, et al. Biophysical characterization of an ensemble of intramolecular i-motifs formed by the human c-MYC NHE III1 P1 promoter mutant sequence. Biophys J. 2010; 99:561–7. [PubMed: 20643075]
28. Xu Y, Sugiyama H. Formation of the G-quadruplex and i-motif structures in retinoblastoma susceptibility genes (Rb). Nucleic Acids Res. 2006; 34:949–54. [PubMed: 16464825]
29. Biffi G, Tannahill D, McCafferty J, Balasubramanian S. Quantitative visualization of DNA G-quadruplex structures in human cells. Nat Chem. 2013; 5:182–6. [PubMed: 23422559]

30. Phan AT, Kuryavyi V, Gaw HY, Patel DJ. Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. Nat Chem Biol. 2005; 1:167–73. [PubMed: 16408022]
31. Fedoroff OY, Rangan A, Chemeris VV, Hurley LH. Cationic porphyrins promote the formation of i-motif DNA and bind peripherally by a nonintercalative mechanism. Biochemistry. 2000; 39:15083–90. [PubMed: 11106486]
32. Han FXG, Wheelhouse RT, Hurley LH. Interactions of TMPyP4 and TMPyP2 with quadruplex DNA. Structural basis for the differential effects on telomerase inhibition. Journal of the American Chemical Society. 1999; 121:3561–3570.
33. Han HY, Langley DR, Rangan A, Hurley LH. Selective interactions of cationic porphyrins with G-quadruplex structures. Journal of the American Chemical Society. 2001; 123:8902–8913. [PubMed: 11552797]
34. Qin Y, Hurley LH. Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. Biochimie. 2008; 90:1149–71. [PubMed: 18355457]
35. Matsushita N, et al. Dynamics of tyrosine hydroxylase promoter activity during midbrain dopaminergic neuron development. J Neurochem. 2002; 82:295–304. [PubMed: 12124430]
36. Thakur S, et al. Regulation of BRCA1 transcription by specific single-stranded DNA binding factors. Mol Cell Biol. 2003; 23:3774–87. [PubMed: 12748281]
37. Kerppola TK, Curran T. Selective DNA bending by a variety of bZIP proteins. Mol Cell Biol. 1993; 13:5479–89. [PubMed: 8355695]
38. Leonard DA, Rajaram N, Kerppola TK. Structural basis of DNA bending and oriented heterodimer binding by the basic leucine zipper domains of Fos and Jun. Proc Natl Acad Sci U S A. 1997; 94:4913–8. [PubMed: 9144164]
39. Nagamoto-Combs K, Piech KM, Best JA, Sun B, Tank AW. Tyrosine hydroxylase gene promoter activity is regulated by both cyclic AMP-responsive element and AP1 sites following calcium influx. Evidence for cyclic amp-responsive element binding protein-independent regulation. J Biol Chem. 1997; 272:6051–8. [PubMed: 9038229]
40. Tinti C, et al. Inducible cAMP early repressor can modulate tyrosine hydroxylase gene expression after stimulation of cAMP synthesis. J Biol Chem. 1996; 271:25375–81. [PubMed: 8810303]
41. Liu N, et al. Unique regulation of immediate early gene and tyrosine hydroxylase expression in the odor-deprived mouse olfactory bulb. J Biol Chem. 1999; 274:3042–7. [PubMed: 9915843]
42. Michelotti EF, Michelotti GA, Aronsohn AI, Levens D. Heterogeneous nuclear ribonucleoprotein K is a transcription factor. Mol Cell Biol. 1996; 16:2350–60. [PubMed: 8628302]
43. Shnyreva M, Schullery DS, Suzuki H, Higaki Y, Bomsztyk K. Interaction of two multifunctional proteins. Heterogeneous nuclear ribonucleoprotein K and Y-box-binding protein. J Biol Chem. 2000; 275:15498–503. [PubMed: 10809782]
44. Goodman RH, Smolik S. CBP/p300 in cell growth, transformation, and development. Genes Dev. 2000; 14:1553–77. [PubMed: 10887150]
45. Brooks TA, Hurley LH. The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. Nat Rev Cancer. 2009; 9:849–61. [PubMed: 19907434]
46. Bomsztyk K, Denisenko O, Ostrowski J. hnRNP K: one protein multiple processes. Bioessays. 2004; 26:629–38. [PubMed: 15170860]
47. Balasubramanian S, Hurley LH, Neidle S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? Nat Rev Drug Discov. 2011; 10:261–75. [PubMed: 21455236]
48. Lam EY, Beraldi D, Tannahill D, Balasubramanian S. G-quadruplex structures are stable and detectable in human genomic DNA. Nat Commun. 2013; 4:1796. [PubMed: 23653208]
49. Jacobs FM, et al. Identification of Dlk1, Ptpru and Klhl1 as novel Nurr1 target genes in meso-diencephalic dopamine neurons. Development. 2009; 136:2363–73. [PubMed: 19515692]
50. Wallen AA, et al. Orphan nuclear receptor Nurr1 is essential for Ret expression in midbrain dopamine neurons and in the brain stem. Mol Cell Neurosci. 2001; 18:649–63. [PubMed: 11749040]
51. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 2004; 32:W273–9. [PubMed: 15215394]

52. Brudno M, et al. LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. Genome Res. 2003; 13:721–31. [PubMed: 12654723]

53. Son JH, et al. Neuroprotection and neuronal differentiation studies using substantia nigra dopaminergic cells derived from transgenic mouse embryos. J Neurosci. 1999; 19:10–20. [PubMed: 9870933]

54. Wood JR, Likhite VS, Loven MA, Nardulli AM. Allosteric modulation of estrogen receptor conformation by different estrogen response elements. Mol Endocrinol. 2001; 15:1114–26. [PubMed: 11435612]

55. Joh TH, Geghman C, Reis D. Immunochemical demonstration of increased accumulation of tyrosine hydroxylase protein in sympathetic ganglia and adrenal medulla elicited by reserpine. Proc Natl Acad Sci U S A. 1973; 70:2767–71. [PubMed: 4147606]

56. Akiba Y, et al. gamma-Aminobutyric acid-mediated regulation of the activity-dependent olfactory bulb dopaminergic phenotype. J Neurosci Res. 2009; 87:2211–21. [PubMed: 19301430]

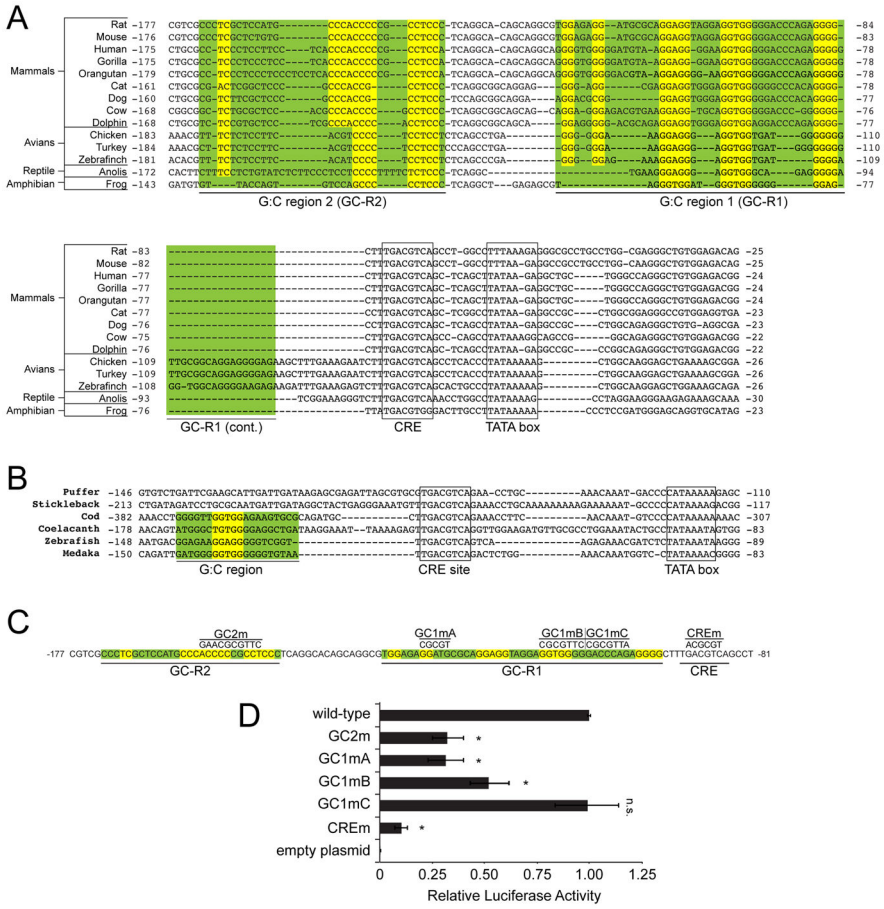

Figure 1.

Identification of highly conserved and functional G:C-rich regions in the vertebrate *Th* proximal promoters. **A** and **B**, alignment of vertebrate *Th* proximal promoter nucleotide sequences. **A**, the *Th* proximal promoter of all tetrapod species examined contains two previously unreported G:C-rich regions (GC-R1 and GC-R2; shaded in green). Within these regions, several short nucleotide motifs are highly conserved (shaded in yellow). Both GC-R1 and GC-R2 are upstream of the highly conserved and well established CRE and TATA box sequences (boxed). **B**, in fish, the CRE and TATA box are also conserved, but a single G:C-rich region was found in only a subset of species. Nucleotide positions are numbered relative to the translation start site. **C** and **D**, functional analysis of G:C-rich regions in the *Th* proximal promoter. **C**, nucleotide sequence of the rat *Th* proximal promoter used for transcription assays. Highly conserved motifs within GC-R1 and GC-R2 are shaded in yellow and mutated positions are shown above the wild-type sequence. The species used for alignment of the proximal promoter sequences were: rat, *Rattus norvegicus*; mouse, *Mus musculus*; guinea pig, *Cavia porcellus*; human, *Homo sapiens*; Gorilla, *Gorilla gorilla*; Orangutan, *Pongo abelii*; cat, *Felis catus*; dog, *Canis familiaris*; cow, *Bos taurus*; dolphin, *Tursiops truncatus*; chicken, *Gallus gallus*; turkey, *Meleagris gallpavo*; zebrafinch, *Taeniopygia guttata*; anole, *Anois carolinensis*; frog, *Xenopus tropicalis*; puffer, *Takifugu rubripes*; Stickleback, *Gasterosteus aculeatus*; Zebrafish, *Danio rerio*; Medaka, *Oryzias latipes*; Cod; *Gadus morhua*; Coelancath, *Latimeria chalumnae*. **D**, luciferase reporter

activity in a *Th*-expressing cell line derived from the mouse olfactory bulb (OB). Mutations overlapping highly conserved nucleotides significantly reduced luciferase reporter activity relative to the wild-type promoter activity (indicated by asterisks; $p<0.01$ by Student's t-test). By contrast, mutations not overlapping highly conserved nucleotides (i.e. GC1mC) did not significantly reduce reporter gene expression (indicated by "n.s."). Luciferase activities are reported as the mean of three independent transfection experiments with error bars representing the standard error of the mean.

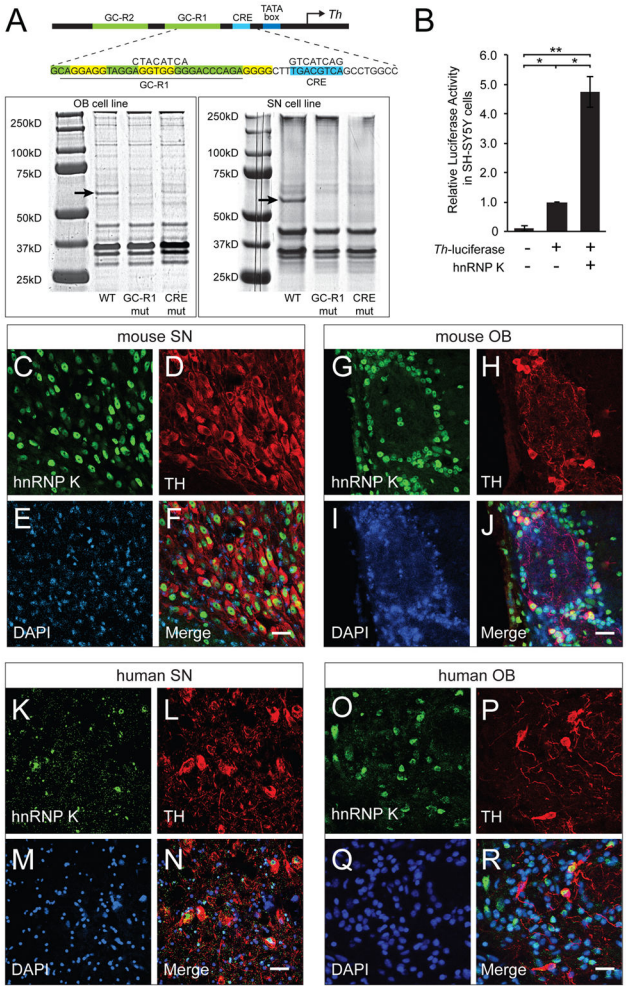

Figure 2.

hnRNP K interacts with the conserved G:C-rich regions in the *Th* proximal promoter and is co-expressed with TH *in vivo*. **A**, protein pull-down assays with nuclear lysates from mouse *Th*-expressing cells derived from the mouse olfactory bulb (OB) and substantia nigra (SN). Forward strand of wild-type and mutant oligonucleotide duplexes derived from GC-R1 are shown with mutant positions above the wild-type sequence. Proteins pulled down with either the wild-type or mutant duplex probes were revealed by coomassie stained SDS-polyacrylamide gels. Mass spectrometry identified the differentially isolated protein (indicated by arrow) as hnRNP K. **B**, transcription assays in human SH-SY5Y cells showed over-expression of hnRNP K could up-regulate *Th* promoter activity. Luciferase activities are reported as the mean of three independent transfection experiments with error bars representing the standard error of the mean. Pairwise comparisons of luciferase activity levels used the Student's t-test (single asterisk indicates $p < 0.01$; double asterisk indicates $p < 0.001$). **C–R**, TH and hnRNP K co-expression in mice and humans. In both the SN (**C–F** and **K–N**) and OB (**G–J** and **O–R**), most TH-containing cells co-express hnRNP K. Bar equals 50 μm.

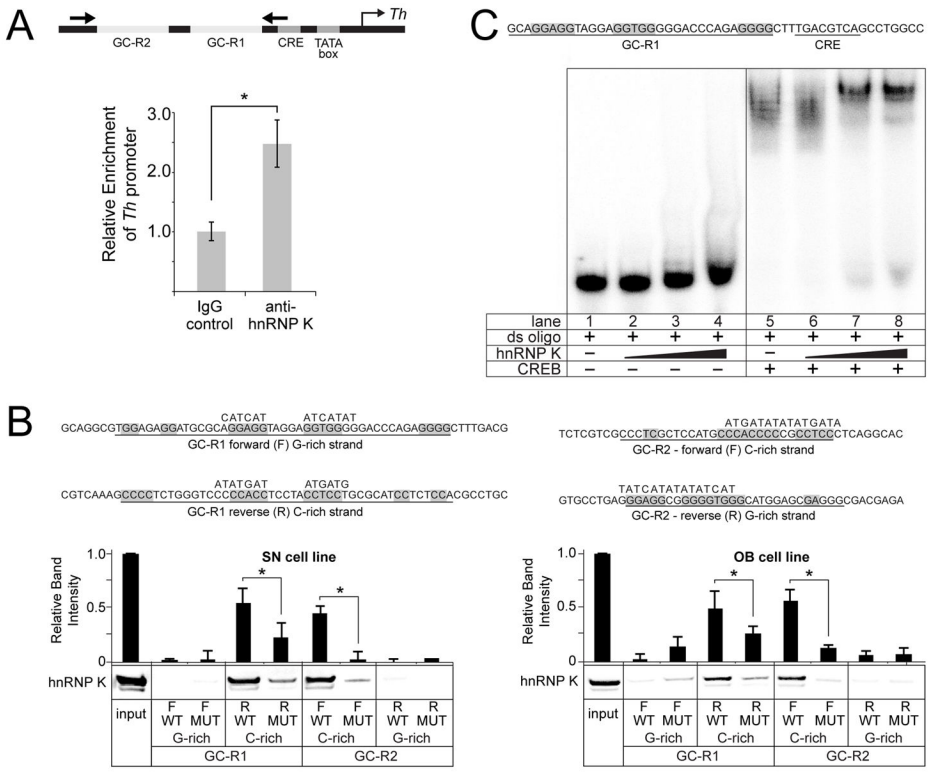

Figure 3.

hnRNP K binding to the *Th* proximal promoter. **A**, chromatin immunoprecipitation assays with human SH-SY5Y cells indicated that hnRNP K directly binds the *Th* proximal promoter. Three independent ChIP experiments were conducted and the mean relative enrichment of the *Th* proximal promoter is reported with error bars representing the standard error of the mean. Statistical significance was assessed using the two-tailed Student's t-test (asterisk indicates $p < 0.01$). Arrows show the position on the *Th* promoter of the primers used in this assay. **B** hnRNP K binds single-stranded sequences from the *Th* promoter. Wild-type and mutant oligonucleotides derived from the GC-R1 and GC-R2 are shown with mutated positions above the wild-type sequence. Highly conserved nucleotides within GC-R1 and GC-R2 are shaded. Western blot analyses showed that wild-type C-rich single strands preferentially pulled down hnRNP K from OB and SN cell line nuclear lysates. Band intensities, shown above the blot, are reported as the mean of duplicate pull-down assays with error bars representing the difference between the mean and the individual trials. Compared to the wild-type strands, mutant C-rich strands significantly reduced the interaction with hnRNP K (asterisk; $p < 0.01$ by Student's t-test). **C**, the duplex oligonucleotide containing GC-R1 (show on top with highly conserved nucleotides shaded) was not bound by hnRNP K in electromobility shift assays (lanes 1–4). By contrast, hnRNP K did super-shift this target sequence when CREB was present (lanes 5–8).

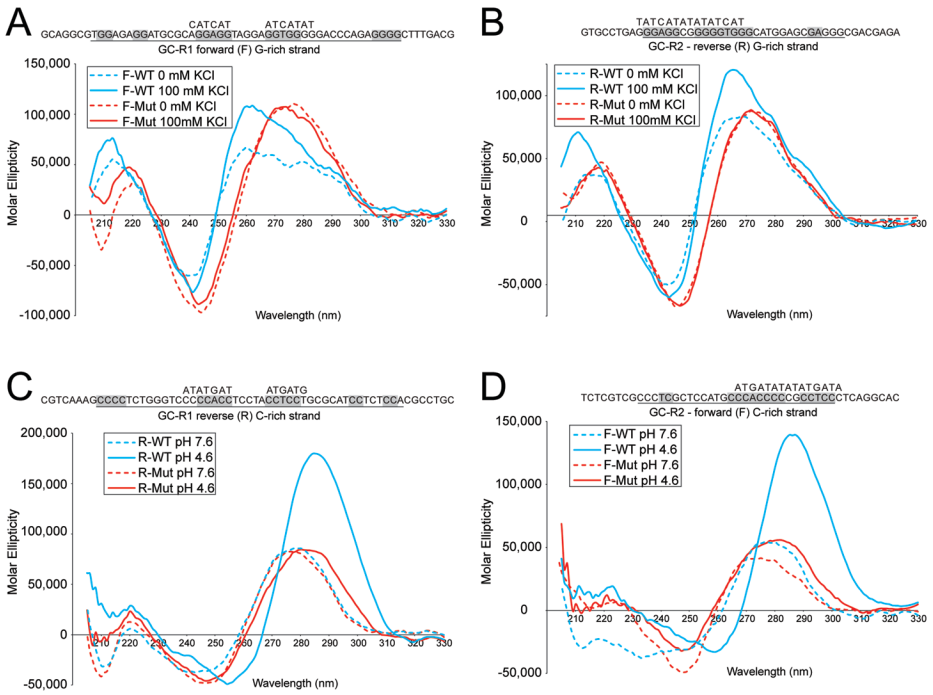

Figure 4.

Circular dichroism spectra for the single DNA strands from G:C-rich regions in the *Th* proximal promoter. **A** and **B**, wild-type G-rich single strands (blue lines) show a KCl concentration-dependent enhancement of spectral features. By contrast, spectral features for mutant sequences (red lines) are not KCl concentration-dependent. **C** and **D**, wild-type C-rich single strands (blue lines) show a pH-dependent wavelength shift and enhancement of signal maxima. By contrast, spectral features for the mutant sequences (red lines) are not pH-dependent. Oligonucleotide sequences used are displayed above their corresponding spectra with highly conserved motifs within these regions shaded and mutant positions are shown above the wild-type sequence.

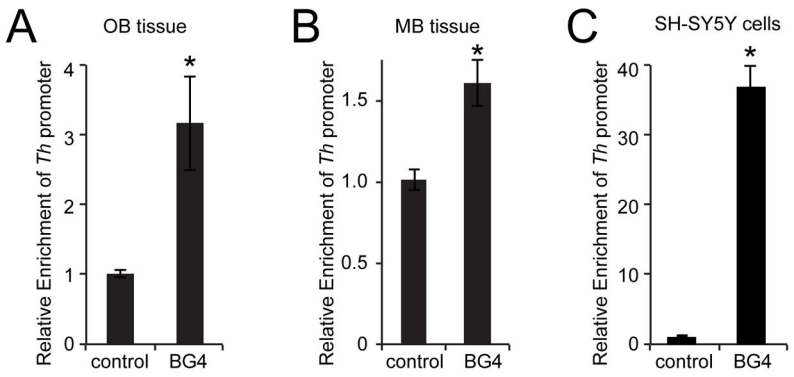

Figure 5.

In vivo and *in vitro* detection of G-quadruplex secondary structures on the *Th* promoter. **A** and **B**, ChIP experiments with the BG4 antibody detected G-quadruplex secondary structures on the *Th* promoter *in vivo* in both the mouse olfactory bulb (OB) and midbrain (MB), respectively. **C**, parallel *in vitro* ChIP studies with human SH-SY5Y cells confirmed the presence of G-quadruplex secondary structures on the *Th* promoter. For ChIP experiments, three independent immunoprecipitations were performed and the mean relative enrichment of the *Th* proximal promoter is reported with error bars representing the standard error of the mean. Statistical significance was assessed using the two-tailed Student's t-test (asterisks indicate $p < 0.01$).

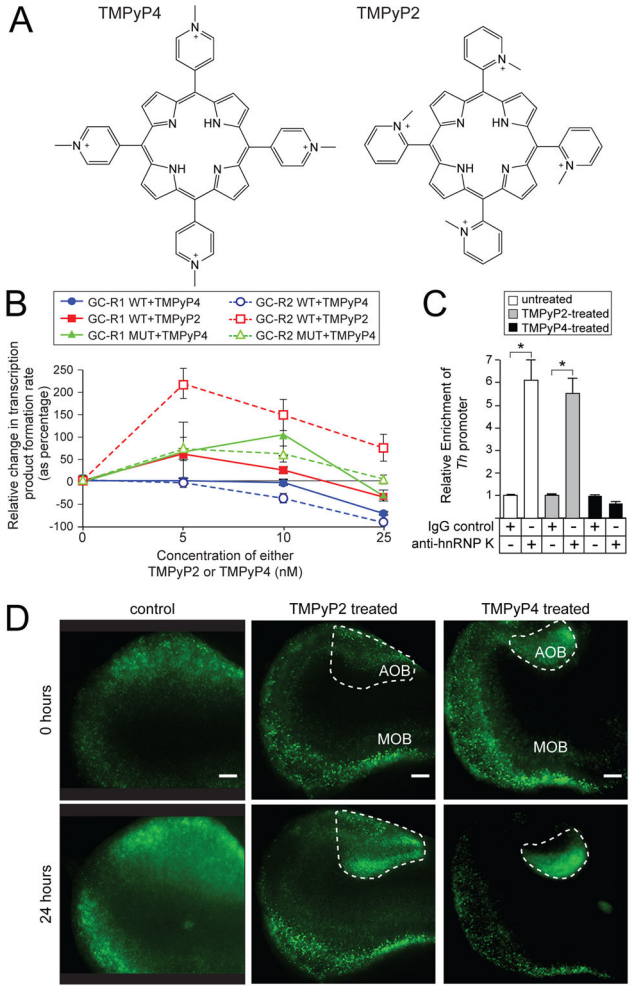

Figure 6.

Small molecule manipulation of DNA secondary structure in the *Th* promoter. **A**, the isomeric porphyrin compounds TMPyP4 and TMPyP2 were used as agents to modulate DNA secondary structure stability. **B**, increasing concentrations of TMPyP4, but not TMPyP2, block transcription of template strands derived from the GC-R1 and GC-R2. By contrast, TMPyP4 was not effective at impeding transcription of template strands with mutations to the G:C-rich regions of either GC-R1 or GC-R2 (template oligonucleotides are the same as the sequences used in Figure 4). The data shown are the mean of three independent trials with error bars representing the standard error of the mean. **C**, ChIP experiments showed that exposure to TMPyP4, but TMPyP2, blocked hnRNP K occupancy on the *Th* promoter in SH-SY5Y cells. For all conditions, three independent ChIP experiments were performed and the mean relative enrichment of the *Th* proximal promoter is reported with error bars representing the standard error of the mean. Statistical significance was assessed using the two-tailed Student's t-test (asterisks indicate $p < 0.01$. **D**, depolarized forebrain organotypic slice cultures show that TMPyP4 and TMPyP2 differentially modulate *Th* promoter activity in dopaminergic neurons of the main olfactory bulb (MOB). After 24 hours of exposure, slices exposed to 100 μM TMPyP4 had substantially lower *Th*-GFP expression in MOB dopaminergic neurons. By contrast, *Th*-GFP

expression in these neurons was not diminished in slices treated with either 100μM TMPyP2 or control conditions. The accessory olfactory bulb (AOB) is distinguished from the MOB by the dotted line. Bar equals 100 μm.

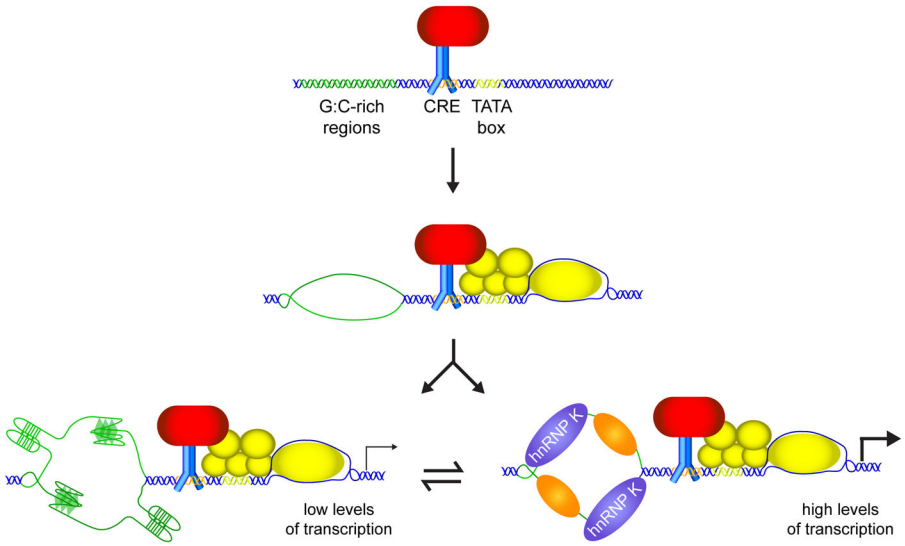

Figure 7.

Proposed regulatory mechanism for hnRNP K and DNA secondary structure on the *Th* proximal promoter. Recruitment of bZIP proteins (in blue) and associated co-activator proteins (in red) to the CRE facilitates assembly of the RNA Polymerase transcription protein complex (in yellow; top panel). Melting of the duplex core promoter DNA by the transcription protein complex generates torsional stress that drives the separation of the duplex DNA in the G:C-rich regions (in green; middle panel). These single DNA strands from the G:C-rich regions either adopt secondary structures, such as G-quadruplexes or i-motifs (bottom left panel), or are bound by proteins such as hnRNP K (bottom right panel). *Th* transcription is diminished when secondary structures are formed, whereas stabilization of the unfolded single strands in the G:C-rich regions enhances transcription.