DNA tetraplex formation in the control region of c-myc

Tomas Simonsson*, Petr Pecinka¹ and Mikael Kubista

Department of Biochemistry, Lundberg Institute, Chalmers University of Technology, Medicinaregatan 9C, SE-413 90 Goteborg, Sweden and ¹Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, CZ-612 65 Brno, Czech Republic

Received December 12, 1997; Revised and Accepted January 20, 1998

ABSTRACT

The c-myc oncogene is one of the most commonly malfunctioning genes in human cancers, and is an attractive target for anti-gene therapy. Although synthetic oligonucleotides designed to silence c-myc expression via one of its major control elements function well in vitro, their mode of action has been indefinite. Here we show that the targeted control element adopts an intrastrand fold-back DNA tetraplex, which requires potassium ions for stability in vitro. We believe formation of the tetraplex is important for c-myc activation in vivo, and propose a transcription initiation mechanism that explains how anti-gene therapy silence c-myc at the molecular level.

INTRODUCTION

The myc family of oncogenes encodes phospho-proteins that activate genes, which force cell growth forward. Normally the human c-myc gene is tightly regulated and alterations in its expression is a key step in the progression of many cancers. It is overexpressed in a variety of malignancies including lymphomas, leukaemias, and lung, cervical, ovarian, breast and gastric cancers (1). The oncogenic properties of c-myc arise from phenomena such as gene amplification (2), chromosomal translocation (3), retroviral transduction (4) and proviral insertion (5). Irrespective of how c-myc regulates oncogenic properties, mechanisms behind deviant c-myc transcription are obscured by the fact that the gene utilises four promoters (6). There is, however, a major control element of the human c-myc oncogene, located (–115) – (–142) bp upstream of the c-myc promoter P1. This element, which corresponds to bases 2186–2212 in the sequence of the human c-myc locus (7), is termed nuclease-hypersensitive element III1 (NHE) (8) and accounts for 75–85% of total c-myc transcription (9,10). The NHE has an unusual strand asymmetry; one strand is an almost perfect homopyrimidine tract and the other is a homopurine tract.

A clue to c-myc activation came from the observation that its transcription could be selectively blocked by an oligonucleotide targeted to a nuclease-hypersensitive element upstream of promoter P1 (11). A synthetic 27-base-long oligonucleotide, complementary to the coding strand of the c-myc gene targeted to the NHE, was first found to inhibit specifically c-myc transcription in HeLa cell extracts. The same oligonucleotide was later shown to inhibit c-myc transcription in live HeLa cells (12) and, when conjugated to acridine, it exhibited a similar effect on other cell lines where c-myc is overexpressed (13). The mechanism suggested has been that the oligonucleotide binds to the NHE, forming a purine–pyrimidine–purine triplex that interferes with the binding of transcription factors (11). However, purine–pyrimidine–purine triplexes require 5–20 mM magnesium, and attempts to substantiate the existence of such triplexes at physiological conditions have failed (14).

Characterrisation of the NHE in vitro has revealed that it is involved in a slow equilibrium between a traditional helix structure and a stable, atypical structure that is sensitive to S1 nuclease (15). It has been proposed that the NHE adopts a tandem H-DNA structure which involves two intramolecular pyrimidine–purine–pyrimidine triplexes, and it was speculated that c-myc is activated by transcription factors that recognise this highly unusual structure (16). Intramolecular pyrimidine–purine–pyrimidine triplexes have been extensively studied in vitro. Although low pH and negative superhelicity favour them, they are essentially independent of ionic composition (17).

The two models both assume that triplex formation in the NHE regulates c-myc transcription. One posits a purine–pyrimidine–purine triplex that requires high magnesium concentrations, and the other postulates pyrimidine–purine–pyrimidine triplexes that require low pH. Consequently, neither is likely to form in live cells and our objective here is to characterise the NHE under physiological conditions.

MATERIALS AND METHODS

Plasmid DNA

The 2776 bp plasmid containing the NHE, pNHE, was constructed by cloning the 105 bp Acii/XmaI restriction fragment of the c-myc promoter sequence into the Acci/XmaI site of pUC18. The clone was transformed into Escherichia coli DH5α cells and spread on agar plates, colonies were grown in LB medium, and subject to flow column purification (Qiagen). The insert was verified on an A.L.F. DNA sequencer (Pharmacia), and by unique

*To whom correspondence should be addressed. Tel: +46 317 733919; Fax: +46 317 733910; Email:tomas@bcbp.chalmers.se
linearisation from Tth111I restriction cleavage. Restriction enzymes were from New England Biolabs, except for AccI and Tth111I, which was from Promega.

**Synthetic oligonucleotides**

The purine-rich 27-base-long oligonucleotide PU27 corresponds to the NHE sequence 5′-TGG GGA GGG TGG GGA GGG TGG GGA AGG-3′. Its complement, the pyrimidine-rich PY27, has the sequence 5′-CCT TCC CCA CCC TCC CCA CCC TCC CCA-3′. Synthetic oligonucleotides were 1 μmol-syntheses from Eurogentec. All oligonucleotides were purified by denaturing gel electrophoresis prior to use. Radiolabelling in the 5′-end was done with [γ-32P]dATP and T4 polynucleotide kinase, and 3′-end radiolabelling was carried out using [α-32P]ddA TP and terminal deoxynucleotidyl transferase. Enzymes were from Promega and 32P-nucleotides were from Amersham.

**Primer extension**

Primer extension reactions were performed using Klenow fragment (Promega) and pNHE as template. The primer 5′-GTCATA GCT GTT TCC TGT GT-3′ was used to extend the homopurine strand of the NHE.

**Chemical probing**

Guanine specific probing was performed for 5 min in 0.5% dimethyl sulphate (Fluka), adenine specific reactions were performed for 15 min in 2% diethyl pyrocarbonate (Aldrich), and thymine specific modification was performed for 15 min in 1 mM osmiumtetroxide (Fluka)-bipyridine (Sigma). All reactions were done at 37°C and followed by piperidine (Sigma) cleavage and extensive lyophilisation. Analyses were performed on 12% polyacrylamide gels (19:1 mono:bis) under standard denaturing conditions.

**Gel electrophoresis**

Acrylamide solutions and ammonium persulfate were from BIORAD, TEMED was from Fluka, TRIS was from Amresco, boric acid was from J.T.Baker, and EDTA was from Merck. Agarose was SeaPlaque from FMC. Lithium-, sodium-, potassium-, rubidium- and cesium-chloride were from Sigma. Native polyacrylamide gel electrophoresis was performed on 0.4 mm-thick 20% polyacrylamide gels (19:1 mono:bis) at 4°C and 8 W for 8 h. Agarose gel electrophoresis was performed in a 2% agarose gel at 4°C and 20 V/cm for 30 min. In all native gel electrophoresis experiments both gel and buffer contained 0.5 TBE and 12.5 mM each of NaCl and KCl.

**RESULTS**

We performed primer extension experiments with the non-coding homopurine strand of pNHE as template (Fig. 1A). The experiment was repeated 12 times which unanimously gave the same result: there was a potassium-dependent DNA synthesis arrest at the location of the NHE. We ascribe the arrest in DNA synthesis to the formation of a potassium-dependent compact structure in the template strand. The structure is exceedingly stable: not even after 90 min was an appreciable fraction of the template elongated beyond the arrest site (Fig. 1B).

To characterise the structure of the NHE in detail, we used the purine-rich PU27, and its complement PY27. When PU27 was pre-incubated in presence of potassium ions, it gave rise to two bands in native polyacrylamide gels (Fig. 2A). One band has the expected electrophoretic mobility of an unstructured oligonucleotide, while the other has a higher mobility. The oligonucleotide of higher mobility was characterised by chemical probing with various base-specific reagents (Fig. 2B and C). The probing was repeated seven times and the pattern was

---

**Figure 1.** (A) Polyacrylamide gel electrophoresis shows that DNA synthesis arrest depends on the presence of potassium ions in primer extension reactions. Primer extension was performed in: lane 1, 100 mM LiCl; lane 2, 100 mM NaCl; lane 3, 80 mM NaCl/20 mM KCl; lane 4, 60 mM NaCl/40 mM KCl; lane 5, 40 mM NaCl/60 mM KCl; lane 6, 20 mM NaCl/80 mM KCl; lane 7, 100 mM KCl; lane 8, 100 mM RbCl; lane 9, 100 mM CsCl. Lanes 10, 11, 12 and 13 are ddNTP reactions specific for bases C, T, G and A, respectively. (B) Polyacrylamide gel electrophoresis shows the time-dependence of the DNA synthesis arrest in the presence of 100 mM potassium. Primer extension was run for: lane 1, 1 min; lane 2, 5 min; lane 3, 15 min; lane 4, 30 min; lane 5, 90 min. Lanes 6, 7, 8 and 9 are ddNTP reactions specific for bases G, A, C and T, respectively. The supercoiled pNHE used as template in both (A) and (B) was pre-incubated at 37°C for 48 h with the same alkali ions that were present during primer extension.
always the same: the four contiguous stretches of guanines, G2–G4, G7–G9, G16–G18 and G21–G23, were protected from dimethyl sulphate modification, which is evidence that they engage in tight stacking and base pairing. The adenines A6 and A15 and the thymines T10 and T19 are partly protected from diethylpyrocarbonate and osmium tetroxide-bipyridine modification, suggesting that they may too participate in tight interactions. The bases G5, G11–G14 and G20 are more susceptible to dimethyl sulphate, which means they should not be involved in tight interactions. The end-most thymine T1 and the protruding 3′ sequence …-A24-A25-G26-G27 display no significant differences to chemical probes when comparing the oligonucleotide of higher mobility and the unstructured form.

The PU27 has the same sequence as the oligonucleotide previously shown to selectively repress c-myc transcription (11). To test how potassium ions influence the interaction of PU27 with the NHE, we pre-incubated pNHE in either 100 mM potassium or sodium, and then mixed it with either radiolabelled PU27 or PY27 (Fig. 3). The PU27, but not its complement PY27, co-migrated with the plasmid when it had been pre-incubated with potassium ions. Neither of the oligonucleotides co-migrated with the plasmid when it had been pre-incubated with sodium ions. There is a prominent sequence specific interaction between the NHE and PU27 that requires potassium ions.

**DISCUSSION**

In our primer extension experiments, the enzyme activity does not depend on the nature of the monovalent cation and we ascribe the arrest in DNA synthesis to the formation of a potassium-dependent compact structure in the template strand. Similar observations have previously been made for related sequences (18).

The only known DNA structure that is both specifically stabilised by potassium ions and exceedingly stable is the DNA tetraplex (19). DNA tetraplexes are built up from guanine tetrads that arise from the association of four guanines in a cyclic Hoogsteen hydrogen-bonding arrangement that involves N1, N7, O6 and N2 of each guanine base. Intrastrand fold-back DNA tetraplexes are formed in regions with four runs of at least two contiguous guanines interspersed by two or more bases; they are
Figure 3. Gel electrophoresis showing that PU27 is specifically retarded when pNHE has been pre-incubated in the presence of potassium ions. Radiolabelled oligonucleotides were mixed with pNHE pre-incubated at 37°C for 48 h in the presence of either sodium or potassium ions. Lane 1, PU27/pNHE pre-incubated in 100 mM NaCl; lane 2, PY27/pNHE pre-incubated in 100 mM NaCl; lane 3, PU27/pNHE pre-incubated in 100 mM KCl, lane 4, PY27/pNHE pre-incubated in 100 mM KCl; lane 5, PU27/pUC18 pre-incubated in 100 mM KCl.

The NHE central region has the necessary motif to form an intranastrand fold-back DNA tetraplex, and its sequence can be written G3N2G3N6G3. The modification pattern from chemical probing of the band of higher electrophoretic mobility is consistent with an intranaststrand fold-back DNA tetraplex structure where G2–G4, G7–G9, G16–G18 and G21–G23 form three planes of guanine tetrads (Fig. 4). These tetrads presumably coordinate two potassium ions. The bases A6 and T19 as well as A15 and T10, we believe, form two intranaststrand AT base pairs that cap the guanine tetrads and add to the stability of the tetraplex structure. It was recently shown that the telomeric repeat of Bombyx mori, 5′-TTAGG-3′, adopts a four-stranded structure in which two guanine tetrads are capped by two (T•A)•A triads in a similar way (25).

It is notable how our finding, that the NHE adopts an intranaststrand fold-back DNA tetraplex structure, is in total agreement with previous characterisations of the NHE. The DNAse I hypersensitivity is confined to the four unpaired guanines G11–G14 in the longer loop of the tetraplex structure (8) (Fig. 4); S1 hypersensitivity is not evenly distributed over the NHE, but is localised around the two bases immediately downstream of the tetraplex structure (15).
actively resolves stable DNA tetraplexes recently was identified (27).

Proteins that bind sequence-specifically to either the homopyrimidine strand or the homopurine strand of the NHE in human c-myc have been identified. Nucleoside diphosphate kinase B, NDPK-B, also known as PuF or NM23-H2 (28), and heterogeneous nuclear ribonucleoprotein K, hnRNP K (29), are c-myc transcription factors that bind to the homopyrimidine strand. Heterogeneous nuclear ribonucleoprotein A1, hnRNP A1 (30), and cellular nucleic acid binding protein, CNBP (31), both bind sequence specifically to the homopurine strand.

NDPK-B has helicase activity and recognises the double-stranded form of the NHE, but binds with higher affinity to the homopyrimidine strand (32). Co-transfection of cells with a plasmid coding for NDPK-B and a plasmid carrying the NHE upstream of a CAT reporter gene revealed that the interaction of NDPK-B with the NHE augments transcription (9). However, it was recently shown that NDPK-B is not involved in transcription activation in the way one would expect for a conventional transcription factor (33).

Despite its name, hnRNP K has a higher affinity for single-stranded DNA than for RNA, and it is virtually oblivious to double-stranded DNA (34). Recently hnRNP K has been shown to interact with the RNA polymerase II transcription machinery in vivo, and its binding motif comprises at least two repeats of the sequence CC(C/T)TCCCCA separated by a spacer (35). The homopyrimidine strand of the NHE has three such repeats.

The hnRNP A1 has been shown to bind to the NHE (30), and it has now been discovered that a rat liver nuclear protein, that is closely homologous to hnRNP A1, binds to an intrastand fold-back DNA tetraplex (36).

CNBP is a zinc finger protein that activates transcription from the NHE in vivo, and has been identified as a major homopurine strand binding protein in HeLa cell nuclear extract using the purine rich strand of the NHE as a probe (31).

How can the transcription factors NDPK-B, hnRNP K, hnRNP A1 and CNBP, in conjunction with tetraplex formation in the NHE, activate c-myc? In Figure 5 we envision a scenario in which NDPK-B unwinds the double-stranded form of the NHE, binds to the homopyrimidine strand. Simultaneously, the homopurine strand forms a tetraplex that is bound by hnRNP A1 and/or CNBP. The exposed homopyrimidine strand is then bound by hnRNP K, which attracts the RNA polymerase II transcription machinery and activates c-myc. This scenario although speculative at this point, suggests that any anti-proliferative effect of the PU27 oligonucleotide originates from efficient competition with c-myc transcription factors that are specific for the homopyrimidine strand. The major role of PU27 in anti-gene therapy is likely to be hybridisation with the homopyrimidine strand of the NHE. When the single-stranded homopyrimidine strand is unavailable, hnRNP K does not recognise the NHE and transcription is blocked since this also prevents interactions between the NHE and the RNA polymerase II transcription machinery.

The sequence motif required to form an intrastand fold-back DNA tetraplex, GxNy1GxNy2GxNy3Gx, is not unique to the c-myc promoter region. It seems to be a recurring theme in the promoter regions of many oncogenes; c-abl (37), c-ets (38,39), c-fes/fps (40), c-fgr (41), c-fos (42), c-jun (43), c-kit (44), c-mos (45), c-myb (46), c-rel (47), c-sis (48), c-src (49), c-yes (50), and the vav proto-oncogene (51) are examples of human cellular oncogenes that exhibit the "intrastand fold-back DNA tetraplex"-motif in their upstream promoter regions. We find it likely that intrastand fold-back DNA tetraplexes constitute a novel transcriptional control element, which provides potent targets for anti-gene therapies.
ACKNOWLEDGEMENTS

Georg Klein, Anders Wennborg, Charlotte Asker, and Emil Palecek are acknowledged for discussion and advice. Part of this work was funded by the Swedish Cancer Society.

REFERENCES