RNA polymerase III initiation on coligo DNA templates containing loops of variable sequence, size and nucleotide chemistry

Joy Patel^a, Lodoé Lama^{a,b}, Niklas A. Hoffmann^c, Jose Cobo^{a,b}, Christoph W. Müller^c and Kevin Ryan^{a,b,*}

^a Department of Chemistry and Biochemistry, The City College of New York, 160 Convent Avenue, New York, New York, USA

^b Biochemistry Doctoral Program, The City University of New York Graduate Center, 365 Fifth Avenue, New York, New York, USA

^c European Molecular Biology Laboratory (EMBL) Structural and Computational Biology Unit, Meyerhofstrasse 1, 67117 Heidelberg, Germany

ABSTRACT Circularized oligonucleotides, or coligos, were previously found to serve as RNA polymerase III (Pol III) templates in vitro and in human tissue culture cells. Here we randomized the 12-nucleotide larger loop (L-loop) of a well characterized coligo and found unexpectedly that in vitro transcription by FLAG-Pol III was not significantly affected. This observation allowed us to test the variable of coligo L-loop size separately from the variable of its sequence. Transcription efficiency increased with L-loop size from 3 to 12 nucleotides of randomized sequence, and the smallest loop forced initiation to move into the stem region. To test further the need for any specific sequence we compared seven nucleotide L-loops composed of random, abasic and abasic-acyclic nucleotides, and all supported transcription by Pol III. Transcription of a series of coligos containing twelve contiguous randomized nucleotides placed at different locations within the coligo structure provided further evidence that the stem-loop junction structure is important for precise initiation. Nearly the same transcript pattern was formed in vitro by Pol III from yeast and human cells. Overall, these experiments support structure, rather than L-loop sequence, as the major determinant of coligo transcription initiation by Pol III.

Key Words: coligo, transcription, RNA polymerase III, yeast RNA polymerase I, II and III, abasic nucleotide

1. Introduction

In principle, genes need not be any larger than the RNA they encode, though the need for a transcriptional promoter and the non-template strand, at a minimum, always make them so. Some non-coding RNA gene products are sufficiently small that their cDNA template strand falls easily within the limits of chemical DNA synthesis, and these comparatively small transcripts are for the most part made in eukaryotes by RNA polymerase III (Pol III) (Dieci et al. 2007). We are exploring circularized oligonucleotides, or coligos, as potential expression vectors for small RNA. In effect, they constitute DNA information templates of minimal size, made from the template strand and only a few nucleotides (nt) larger than the RNA they encode. In previous work we found that pre-miRNA-encoding coligos can template the synthesis of surprisingly well-defined Pol III transcripts in human cells, cell extracts (Seidl et al. 2013; Lama et al. 2014) and by FLAG-tagged Pol III (Lama et al. 2014), and offer several advantages over their linear counterparts. In cells and extracts, Pols I and II did not appear to contribute to coligo transcription. Pre-miRNA-encoding coligos have one smaller single-stranded (ss) loop and one larger ss loop (L-Loop) flanking an imperfectly base-paired stem

corresponding to the pre-miRNA stem. Previous work has suggested that the coligo stem-loop structure at the L-loop is important for initiation, but did not rule out a significant contribution by the L-loop sequence. In this report, we use loops composed of randomized and abasic nucleotides to rule out any major contribution of loop sequence to initiation, apart from its contribution to secondary structure and circular topology. We also extend our findings to yeast Pol III, which is more commonly used for in vitro biochemical studies.

2. Materials and methods

2.1 Coligo preparation

5' phosphorylated synthetic linear DNA Ultramer® oligonucleotides were purchased from Integrated DNA Technologies (IDT). In the case of randomized nucleotides, “N,” IDT’s machine mix randomization was used (coligos N7, N12, N15, JctN12, AN12, BN12, CN12, SLN5). In coligo precursor Ab, which is the same size as N7, all seven nucleotides in the 7 nt loop were replaced with 1',2'-dideoxyribose (IDT dSpacers) to introduce consecutive abasic sites. Similarly, in the linear precursor to coligo C3, all seven nucleotides in the larger loop were replaced with consecutive IDT C3 Spacers, which are simple n-propyl groups between phosphates. Coligo secondary structures were predicted using the mfold program (Zuker 2003). These DNA structures have not been experimentally verified, for example by DNase footprinting, but the large number of possible base-pairs in the stem and the circular topology of the coligos make these predictions highly probable. All linear precursors were circularized using TS2126 RNA ligase 1, gel purified on a denaturing polyacrylamide electrophoresis gel (DPAGE) and verified to be circular using Nuclease S1 nicking (Supplementary Figure S1), as described previously (Seidl and Ryan 2011).

2.2 HEK293T cell extract preparation and FLAG-RNA Pol III immunoprecipitation

Whole cell extract (WCE) from plated HEK293T cells was prepared exactly as previously described (Lama et al. 2014). The human RNA polymerase III complex was immunoprecipitated from WCE made from plated HEK293/POLR3F cells (FLAG tag is on the N-terminus of the POLR3F subunit, (Chiu et al. 2009)) using the FLAG® Immunoprecipitation Kit (FLAGIPT1, Sigma-Aldrich) as previously described (Lama et al. 2014). A 10-fold scaled up immunoprecipitation was separated from the 3XFLAG elution peptide by 12% SDS-

PAGE and subjected to trypsin LC-MS analysis by the Rockefeller University Proteomics Facility (Supplementary Fig. S2.) All but the smallest Pol III subunit, which was likely removed with the FLAG peptide during pre-LC-MS electrophoresis, were identified. Yeast RNA polymerase II was a gift from P. Robinson and R. Kornberg, Stanford University. Yeast RNA Polymerase III was prepared as previously described (Hoffmann et al. 2015). Yeast RNA polymerase I was prepared as previously described (Fernandez-Tornero et al. 2013).

2.3 In vitro transcription

Transcription reactions were carried out in a 20 μ l volume as previously described (Lama et al. 2014). Reactions using whole cell extract as the Pol III source contained 4 μ l extract (from a 1400 μ l extract preparation resulting from three 10 cm plates grown to ~90% confluency). Reactions using immunoprecipitated FLAG-Pol III contained 4 μ l of FLAG antibody beads (i.e. 8 μ l of a 50% suspension) with bound enzyme, as described (Lama et al. 2014). For transcription reactions using bead-bound RNAP III complex, the amount of enzyme used per reaction was equivalent to 2% of the total input WCE made from three 10 cm plates for immunoprecipitation. Reactions using yeast RNA polymerases were diluted to the final concentrations indicated in the figures. Transcription reactions also contained: 8 U RNasin® RNase inhibitor (N2515, Promega), 0.6 mM each ATP, GTP, CTP, 0.2 mM UTP, ~2 μ Ci [α -³²P]-UTP, 40 mM Tris-HCl pH 7.9 @ 25°C, 6 mM MgCl₂, 10 mM DTT, 2 mM spermidine, and 100 nM DNA coligo or other template. A 10-fold scale up of a typical transcription reaction was performed for the library preparations shown in Fig. 3 so that the RNA could be visualized using Kodak film in a two hour exposure, and excised. Transcription reactions were incubated for 90 min at 37 °C followed by extraction with TRIzol® reagent (Invitrogen, 15596-026) and precipitation following manufacturer's instructions. Precipitated RNA was separated by 10% DPAGE, and visualized by exposing the dried gel to a Molecular Dynamics Phosphorimager Screen. Image quantification was performed using MD ImageQuant software.

2.4 Library preparation for high throughput sequencing of transcripts from coligos N7, N12, and N15.

Monomer transcript bands from 10-fold scaled up transcription reactions, using immunoprecipitated RNAP III and indicated in Fig. 3A, were used for small RNA-seq library preparation. The library preparation method, which we refer to as Coligo-seq, is similar to the method used in ribosome footprint profiling (Ingolia et

al. 2009), and was adapted to the Illumina sequencing platform. The libraries were sequenced on Illumina Miseq (110 cycles) at the Personalized Genomic Medicine Laboratory, Department of Pathology and Cell Biology, Columbia University Medical Center. The Coligo-seq library preparation method will be reported in detail elsewhere.

2.5 Sequencing data processing and analysis

Bioinformatic analysis was done using freely available bioinformatics tools on the www.usegalaxy.org website. Reads obtained for each library were divided into different pools based on their lengths. Reads within a given length were then trimmed to select the first 20 bases from each read which were then quality filtered to remove poor quality reads. The remaining good-quality reads (phred quality score equal to or greater than 20 for each base) were “collapsed” to find the identical constant sequences in the coligo stem, and their count. Each unique read was then manually aligned with respect to their initiation site relative to stem-loop junction while ignoring the actual sequence from the randomized loop region. Finally, the aligned reads from each length category were added up, again focusing on the initiation position preference, with respect to the constant stem-region sequence, instead of L-loop nucleotide identity preference.

3. Results and Discussion

3.1 The coligo larger loop sequence is not a major determinant of Pol III initiation

In order to develop coligos as small RNA expression vectors it is necessary to understand how their primary and secondary structures relate to Pol III initiation and termination. If coligos can be systematically designed to yield Pol III transcripts with predictable and well-defined 5' and 3' ends, then the RNA they engender may enter endogenous RNA processing pathways and result in miRNA mimics and siRNA. One variable that we have not systematically addressed is the sequence in the larger ss loop, the L-loop. In a typical coligo, most initiation and termination seems to take place in or near this loop. For example, coligo 122 (89 nt, Fig. 1A), based on the pre-miRNA of human miR-122, is transcribed by Pol III mainly into a transcript of 83 nt (Seidl et al. 2013; Lama et al. 2014). We replaced 122's L-loop sequence (...ATAGTTTAGACA...) with a 12 nt randomized sequence where each nucleotide consisted of a mixture of approximately 25% A,C,T and G,

represented by “N.” As shown in Fig. 1B, in vitro transcription by FLAG-Pol III produced similarly sized main “monomer,” i.e. single circumtranscription, products from 122 and N12. The 122 monomer transcripts have been sequenced previously using 5’ RACE, and the sequence of the dominant transcript is depicted by the transcription start site arrow and stop indicator mapped to the coligo secondary structure (Fig. 1A). (Small amounts of larger tandem multimer transcripts are also typical and indicated by the spiral arrow to the right in Fig. 1B; note also that abortive transcripts are more prominent in the FLAG-Pol III experiments, most likely because the nucleases that degrade them in extract have been removed.) Since the loop sequence of 122 is only one of 4¹² sequences in N12, and its concentration in this mixture would be too low to be transcribed (Seidl et al. 2013), we take this result to mean that the 122 larger loop sequence is not required to form its main transcripts.

3.2 Larger loop size and Pol III initiation site and efficiency

This unexpected result made us aware that we could evaluate the L-loop size as a factor contributing to transcription efficiency. When making point mutations and other small changes in the L-loop, it is possible to inadvertently produce alternative secondary structures whose stability might reduce the abundance of the coligo stem loop structure. The use of randomized nucleotides eliminates this possibility since such alternative structures will comprise a very small percent of the total structures, which will be dominated by the constant stem sequence region. Thus, by using randomized nucleotides we can vary the coligo 122 L-loop size and assess the transcripts made by Pol III. Accordingly, transcription of coligo N12 was compared to N3 (3 consecutive random nucleotides in place of 122’s L-loop) and similarly, N7 and N15, in HEK293T whole cell extract (Fig. 2A) and FLAG-Pol III from HEK293 cells (Fig. 2B). As the L-loop size was increased, we found a corresponding increase in monomer transcript abundance formed by both Pol III sources. However, the size of the main transcripts changed as well, increasing with loop size especially as the loop was increased from 3 to 7 nt. By N7, most of the transcripts made were the same size as those made from coligos 122 and N12, though the shorter transcripts (about 70 nt compared to 83 nt) still accounted for 5-10% of the monomer transcripts from N7.

To understand why the size of the monomer transcripts were increasing with randomized loop size we excised the monomer transcripts made by FLAG-Pol III from coligos N7, N12 and N15 (Fig. 3A) and determined their transcription start sites using a modified RNA-seq protocol tailored to coligo transcripts (see

Materials and Methods). Most transcript start sites mapped to the predicted stem-loop junction, especially to the first few single-stranded nucleotides of the N loop (i.e. -1, -2 etc; in Figs. 1A and 3B the junction is assigned to 0, the stem nucleotides are positive and increase in the direction of transcription, loop positions are negative). However, only the smallest, N7, showed a significant percentage of transcripts that had initiated at the +6 position, an unpaired C common to all three coligo 122-based templates (see Fig. 1A). Assuming common termination sites in this series of coligos, the smaller transcripts observed in N7, and predominating in N3, which were not abundant enough for our sequencing protocol, likely started in the stem at +6. Thus, when presented with a very small L-loop, Pol III appears to have an increased preference for moving the transcription start site into the stem (in addition to an overall reduced tendency to initiate). This observation may mean that Pol III has to make its own L-loop by melting a short segment of helix when a sufficiently large pre-formed loop is not available. Presumably, a stretch of single-stranded nucleotides is needed to bind the polymerase as a transcription bubble mimic before initiation is triggered. We speculate that the need for a coligo stem melting event, spontaneous or Pol III-mediated, may lead to lower amounts of monomer transcript (Fig. 2AB, lanes 5, 5'). We note that the excised bands did not contain RNA initiating at the loop at the other end of the stem (the AATTG loop), though small percentages of transcripts beginning at the large internal loop starting at +20, were found: N7, 1.5%; N12, 1.6%; N15, none.

3.3 Pol III transcription from coligos with abasic nucleotides in the larger loop

The N12 randomized loop allowed for almost as much transcription as observed for the original coligo 122 larger loop. We decided to push this line of inquiry further by removing the sequence-defining parts of the nucleotides—the nucleobases—together. Perhaps the enzyme is indifferent to the loop sequence but needs natural nucleobases to sustain an interaction with the coligo as a transcription bubble mimic. We substituted 7 consecutive abasic nucleotides for the 7 randomized nucleotides in the N7 loop (Fig. 4). (N7, rather than the more efficient N12, was used as the reference to minimize the cost of these custom nucleotides). We also stripped the N7 nucleotides down to their minimal 3-carbon propyl C3 linker. Both non-natural loops supported the formation of monomer transcripts by Pol III (lanes 5 and 7), though slightly less RNA was made compared to N7 (lane 3). Interestingly, the RNA made was distinctly smaller, an observation we attribute to the enzyme's need to both initiate and terminate in regions of natural nucleotide structure, which in these cases started in the

stem. In addition, and as expected, Pol III could not transcribe through the non-natural loops, and the amount of tandem multimer transcripts was dramatically reduced (lanes 5 and 7 compared to lanes 1 and 3). These results are consistent with the idea that at least part of the function of the larger coligo loop is merely to form a physical link across the 3' and 5' end of the coligo stem helix, and maintain the integrity of the stem-loop structure. Mixing natural and non-natural nucleotides in the larger loop may enable us to more precisely control Pol III initiation and termination.

3.4 Coligo secondary structure requirements by location

Since the N12 randomized sequence had little effect on the transcription templating ability of coligo 122, we wondered if we could use it to probe other portions of the coligo anatomy to learn where in the coligo its secondary structure was most required. We did this by moving the N12 sequence counterclockwise around coligo 122, as shown in Fig. 5. For example, sliding the N12 sequence six positions counterclockwise places the randomized sequence half over the stem and half over the loop (JctN12). Moving it six more positions results in coligo AN12. Coligos BN12 and CN12 were designed by continuing this strategy. In each case, when the N12 sequence is moved the original coligo 122 nucleotides occupy all other positions. In coligo SLN5 the smaller loop was randomized but kept at 5 nt. To avoid circularizing SLN5 in the randomized sequence, which might result in ligation sequence bias, we moved the circularization site to the other end. This produced a linear control template resembling a tailed template (Dedrick and Chamberlin 1985; Bardeleben et al. 1994), perhaps explaining why this linear template worked better than most other linear control templates (Fig. 5B, lane 13). The results of Pol III action on these coligos are shown in Fig. 5B. As the randomized sequence was moved out of the loop, through the junction and into the stem, the transcripts became less precise and less abundant, until in AN12 very little monomer transcripts were made. Coligo AN12 likely forms one very large (36 nt) ss loop and an abbreviated stem region, quite different from the normal coligo secondary structure. In this case alone, a large amount of high molecular weight rolling circle (Daubendiek et al. 1995) transcripts were found near the gel loading site. Coligos BN12 and CN12 formed no distinct RNA bands at all. This experiment revealed the high degree to which transcription relies on the stem-loop structure of coligos such as these, which encode pre-miRNA-like transcripts, and also further implicates the stem-loop junction as a critical structure within the coligo necessary for discrete transcription initiation.

3.5 Coligo transcription by yeast compared to human Pol III in vitro

From in vitro transcription to structural studies, most biochemical research into Pol III over the years has focused on the yeast enzyme. In case coligos might prove to be useful tools in Pol III research, we compared coligo transcription by human and yeast Pol III (Fig. 6A). Although the yeast enzyme (Hoffmann et al. 2015) was more highly purified than our FLAG-tagged human enzyme, both enzymes gave rise to nearly identical patterns of transcripts from coligo 122, though the yeast enzyme produced more abortive transcripts (not shown, but see also Fig. 6B). The yeast enzyme appeared to produce more monomer transcript as well, but this difference might be attributable to small inaccuracies in the protein concentration estimate. The important finding is that the yeast and human enzymes produced the same transcripts, indicating conservation of the coligo-transcribing ability.

The three eukaryotic RNA polymerases are specific for their own promoters, though some exceptions have been found where more than one polymerase can use the same promoter (James Faresse et al. 2012; Duttke 2014; Padilla-Mejia et al. 2015). Using specific enzyme inhibitors in human cells and extracts (Seidl et al. 2013), we previously found evidence of coligo transcription exclusively by Pol III. Nevertheless, we took the opportunity to compare all three yeast RNA polymerases side by side at the same concentration in vitro (Fig. 6B). In this promoter-independent context, each polymerase showed clear evidence of transcription, but each gave rise to a different pattern of transcripts from the same coligo, likely indicating different preferred initiation or termination regions in coligo 122. At 25 nM concentration, Pol I was the most active and, like Pol III, gave a noticeably repeating pattern of tandem multimer transcripts. Pol II and especially Pol I showed evidence of poor termination after one circumtranscription in comparison to Pol III, though coligo 122 lacks a run of 4 or more A's, the canonical Pol III termination signal in the context of normal ds templates. Overall, of the four enzymes tested, human Pol III gave the fewest transcripts outside of the monomer transcript, and yeast Pals I and III showed the most initiations per coligo.

3.6 A speculative structural model of Pol III initiation at the stem-loop junction

Why does Pol III tend to initiate at the coligo stem-loop junction? We have previously speculated that the L-loop mimics a structural intermediate on the pathway to normal promoter-directed transcription initiation or, otherwise stated, it serves as a transcription bubble mimic (Seidl et al. 2013). Based on the recent Pol III

cryo-EM structure we can further speculate that during initiation the coligo stem-loop junction is positioned where the transcription bubble fork is found in transcribing Pol III (Hoffmann et al. 2015). As depicted schematically in Fig. 7, this orientation of the coligo would simultaneously enable downstream anchorage of the coligo duplex region in the Pol III cleft and upstream positioning of a limited number of stem-loop junction nucleotides at the active site.

4. Conclusions

The coligo L-loop appears to be where Pol III initiates and terminates transcription of pre-miRNA-encoding coligo templates. We find here that its sequence is not critical, but its integrity and size influence the transcription initiation site and efficiency. Our studies show that non-natural nucleotides can be placed in this loop without loss of transcription, a finding that may enable us to evaluate other non-natural nucleotides for their potential to improve the precision of initiation and termination, while also reducing the number of natural single-stranded nucleotides vulnerable to endonuclease degradation.

Acknowledgements J.P. is a 2016 City University of New York Jonas E. Salk Scholar, and in addition gratefully acknowledges support from the National Science Foundation REU 1156512 CCNY Biology Department summer 2015 program. J.C. was the recipient of a fellowship award from the U.S. Department of Education Graduate Assistance in Areas of National Need (GAANN) Program in Molecular Biophysics and Biomaterials at The City College of New York (PA200A120211). K.R. acknowledges support from the National Institutes of Health grants 5SC1GM083754 and 1R21GM073944, and the Army Research Office under grant number W911NF-13-1-0148. C.W.M. acknowledges support by an ERC Advanced Grant (ERC-2013-AdG340964-POL1PIC). Additional infrastructure support at the City College of New York was provided by grant 5G12RR003060 from the National Center for Research Resources and grant 8G12MD7603 from the National Institute on Minority Health and Health Disparities. We thank Peter Nagy, Columbia University Medical Center, for Illumina platform sequencing, and P. Robinson and R. Kornberg for RNA polymerase II.

Appendix A. Supplementary data

Supplementary Figure S1 Evidence for coligo circularization using limited S1 Nuclease digestion.

Supplementary Figure S2 LC-MS analysis of FLAG-Pol III immunoprecipitate

References

- Bardeleben C, Kassavetis GA, Geiduschek EP. 1994. Encounters of *Saccharomyces cerevisiae* RNA polymerase III with its transcription factors during RNA chain elongation. *J Mol Biol* **235**: 1193-1205.
- Chiu YH, Macmillan JB, Chen ZJ. 2009. RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. *Cell* **138**: 576-591.

- Daubendiek SL, Ryan K, Kool ET. 1995. Rolling-Circle RNA Synthesis: Circular Oligonucleotides as Efficient Substrates for T7 RNA Polymerase. *J Amer Chem Soc* **117**: 7818-7819.
- Dedrick RL, Chamberlin MJ. 1985. Studies on transcription of 3'-extended templates by mammalian RNA polymerase II. Parameters that affect the initiation and elongation reactions. *Biochemistry* **24**: 2245-2253.
- Dieci G, Fiorino G, Castelnuovo M, Teichmann M, Pagano A. 2007. The expanding RNA polymerase III transcriptome. *Trends Genet* **23**: 614-622.
- Duttke SH. 2014. RNA polymerase III accurately initiates transcription from RNA polymerase II promoters in vitro. *J Biol Chem* **289**: 20396-20404.
- Fernandez-Tornero C, Moreno-Morcillo M, Rashid UJ, Taylor NM, Ruiz FM, Gruene T, Legrand P, Steuerwald U, Muller CW. 2013. Crystal structure of the 14-subunit RNA polymerase I. *Nature* **502**: 644-649.
- Hoffmann NA, Jakobi AJ, Moreno-Morcillo M, Glatt S, Kosinski J, Hagen WJ, Sachse C, Muller CW. 2015. Molecular structures of unbound and transcribing RNA polymerase III. *Nature* **528**: 231-236.
- Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**: 218-223.
- James Faresse N, Canella D, Praz V, Michaud J, Romascano D, Hernandez N. 2012. Genomic study of RNA polymerase II and III SNAPc-bound promoters reveals a gene transcribed by both enzymes and a broad use of common activators. *PLoS genetics* **8**: e1003028.
- Lama L, Seidl CI, Ryan K. 2014. New insights into the promoterless transcription of DNA coligo templates by RNA polymerase III. *Transcription* **5**: e27913.
- Padilla-Mejia NE, Florencio-Martinez LE, Moreno-Campos R, Vizuet-de-Rueda JC, Cevallos AM, Hernandez-Rivas R, Manning-Cela R, Martinez-Calvillo S. 2015. The selenocysteine tRNA gene in leishmania major is transcribed by both RNA polymerase II and RNA polymerase III. *Eukaryot Cell* **14**: 216-227.
- Seidl CI, Lama L, Ryan K. 2013. Circularized synthetic oligodeoxynucleotides serve as promoterless RNA polymerase III templates for small RNA generation in human cells. *Nucleic Acids Res* **41**: 2552-2564.
- Seidl CI, Ryan K. 2011. Circular single-stranded synthetic DNA delivery vectors for microRNA. *PLoS One* **6**: e16925.
- Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406-3415.

Figure Captions

Fig. 1. Randomizing the coligo 122 larger loop does not adversely affect Pol III transcription. (A) The sequence and predicted secondary structures of coligos 122 (left) and N12 (right). N indicates randomized nucleotides in the larger loop (LL); the sequences are otherwise identical. The arrow on 122 denotes transcription initiation, and the “Stop” indicates apparent transcription termination, from an earlier sequencing study. SL, smaller loop. Small black triangle marks the ligation site used for circularization. The +1/-1 border indicates the 3' side of the stem-larger loop junction. (B) In vitro transcription of coligos 122 and N12 by HEK293T whole cell extract

(lanes 2-4) or by FLAG-Pol III precipitated from HEK293 cells (lanes 5,6). Bands visualized by [α - 32 P]-UTP incorporation. Curled arrow marks monomer transcripts formed by a single circumtranscription. The spiral arrow marks larger tandem multimer transcripts formed by multiple passes around the coligo. Multimers contain multiples of incorporated UTP and are thus overrepresented.

Fig. 2. Coligo in vitro Pol III transcription efficiency and RNA size as a function of increasing larger loop size. In vitro transcription of four coligos, derived from coligo 122 (see Fig. 1A), where the larger loop was replaced by 3, 7, 12, or 15 consecutive randomized nucleotides. L, linear template; C, circular template. (A) In vitro transcription in HEK293T whole cell extract. (B) In vitro transcription by FLAG-Pol III from HEK293 cells. Curled arrows are described in Fig. 1 legend.

Fig. 3. RNA Pol III transcription initiation site preference as a function of larger loop size. (A) In vitro transcription reactions by FLAG-Pol III were scaled up and the monomer transcript bands indicated by the brackets were excised for RNA-seq analysis. (B) Summary of preferred transcription initiation sites mapped to coligos N7, N12 and N15, where +1 indicates the beginning of the base-paired stem, and -1 indicates the first randomized nucleotide in the larger loop (see Fig. 1A).

Fig. 4. Pol III transcription of coligo N7 and variants where the seven randomized loop nucleotides of N7 are replaced by seven consecutive abasic nucleotides or seven C3 propyl linkers. The chemical structures show the repeating unit of the loop, phosphates are not shown. Curled arrows are described in Fig. 1 legend.

Fig. 5. Randomized nucleotides are tolerated in the coligo terminal loops, but not in the stem. By systematically moving the N12 sequence around coligo 122, the regions that should remain structured for transcription to take place are revealed. (A) Sequences of the coligos harboring randomized 12 nucleotide stretches at different positions in the coligo 122 secondary structure. Implied base-pairs reflect the 122 structure while most N's are probably not base-paired in the depicted library member coligo. (B) In vitro transcription by FLAG-Pol III using

coligo 122 and the variants shown in panel (A). L, linear; C, circular. Curled arrows are described in Fig. 1 legend.

Fig. 6. A comparison of coligo transcription by yeast RNA polymerase I, II, III and human Pol III. (A) Human and yeast Pol III enzymes were compared side by side for their ability to transcribe coligo 122 in vitro. FLAG-Pol III was from HEK293 cells. (B) A comparison of in vitro transcription by yeast Pols I, II, III and HEK293 Pol III at 25 nM each, using coligo 122. hs, human HEK293 enzyme; sc, *S. cerevesiae* enzymes.

Fig. 7. Schematic model of a coligo-Pol III initiation complex. This speculative model depicts a possible orientation for the coligo positioned in the cleft where the DNA template normally binds, as found in the yeast Pol III structure (Hoffmann et al. 2015).

A

B

Larger loop nucleotide (x7):

A

B

A

B