UPTEC X 10 026
Examensarbete 30 hp
Oktober 2010

# Rolling circle transcription on smallest size double stranded DNA minicircles

Anders Kristoffersson

UPPSALA UNIVERSITET

# Molecular Biotechnology Programme

Uppsala University School of Engineering

| **UPTEC X 10 026** | **Date of issue 2010-11** |
|---|---|


Author

**Anders Kristoffersson**


Title (English)

# Rolling circle transcription on smallest size double stranded DNA minicircles

Title (Swedish)


Abstract

RNA polymerase T7 is utilized as a component of motor complexes in DNA nanotechnology due to its high promotor specificity, the lack of external transcription factors and its very high processivity, but there is no experience of its application on small double stranded DNA circles. Circular templates from 210 to 126 bp in circumference sharing common promoter and termination motifs were synthesized and transcription was monitored at end point on gel and in real time with a 2' O methyl RNA molecular beacon. RNAP T7 was found to be able to utilize circular dsDNA templates down to 126 bp with moderate impact on transcription rate for saturated systems and rolling circle transcription products were evident with denaturing PAGE gel electrophoresis for templates down to 167 bp.



Keywords

RNA polymerase, RNAP T7, DNA nanotechnology, molecular motor, real time monitoring of transcription, 2' O methyl RNA, molecular beacon, minicircle


Supervisors

**Prof. Michael Famulok**
**Universität Bonn**

Scientific reviewer

**Prof. Gerhart Wagner**
**Uppsala Universitet**

| Project name | Sponsors |
|---|---|
| Language **English** | Security None |
| **ISSN 1401-2138** | Classification |
| Supplementary bibliographical information | Pages **64** |


**Biology Education Centre** Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 471 4687

# Rolling circle transcription on smallest size double stranded DNA minicircles

*Anders Kristoffersson*

## Populärvetenskaplig sammanfattning

Nanoteknik är ett snabbt växande forskningsfält med tillämpningar i vår vardag. Exempel på användningsområden nära oss inbegriper målarfärg, självrengörande fönster och solceller men kanske främst de integrerade kretsar som gör våra datorer och mobiltelefoner möjliga. I detta arbete utreder jag en möjlig komponent till en nanomotor, små DNA cirklar längs vilka ett enzym, ett RNA polymeras rör sig och ger rotationskraft. En sådan motor skulle fungera genom att RNA polymeraset vid transkriptionen rör sig längs med och läser av DNA molekylen. En kopia av DNA koden bildas i form av en RNA molekyl genom hopsättning, eller polymerisering, av fria RNA komponenter. Vid polymeriseringen frigörs energi som driver enzymets rörelse längs den kopierade DNA molekylen. Enzymets rörelse relativt DNA molekylen kan då utnyttjas i en nanomotor.

Medan det finns exempel på tidigare användning av RNA polymeras som molekylär motor för att dra last längst DNA molekyler har detta varit för linjära DNA, ingen har hittills tittat på transkription från DNA-cirklar i nanostorlek. I detta arbete har jag konstruerat en storleksserie av små DNA cirklar och genom att studera storleken av de bildade transkriptionsprodukterna visat att stabil transkription från dessa små cirklar är möjligt. Genom att mäta hastigheten för bildandet av RNA-produkten vid transkription från de olika cirklarna relativt ett linjärt DNA konstateras att påverkan av cirkelarisering på transkriptionseffektiviteten är måttlig. Dessa fynd visar att stabil transkription är möjlig även för mycket små cirkulära DNA molekyler och pekar ut vägen mot framtida DNA-nanomotorer.

**Examensarbete 30 HP**

**Civilingenjörprogrammet Molekylär Bioteknik**
**Uppsala Universitet november 2010**

# Table of Contents

# Background

The purpose of this work has been to investigate the action of the T7 RNA polymerase on small DNA circles in the size down to a hundred base pairs for the possible future application in a DNA nano motor. While it is previously known that rolling circle transcription from small DNA circles is possible this has been for partly single stranded circles which sacrifice some of the advantageous properties of using the DNA as a structural material. In order to study the transcription from DNA nano circles a number of questions about model design and detection systems have been solved, the answers of which are presented here.

## DNA nanotechnology

DNA nanotechnology is the application of DNA as a material to construct nano scale patterns or structures as opposed the information carrying role the molecule fills in nature. DNA has diverse potential applications such as molecular computing, as a scaffold serving as structural guides for other molecules or as programmable assembly chains. I will here describe the principles of using DNA as a structural material; the functions of a few noteworthy DNA nano machines as well as a DNA nanostructure incorporating the concept of a mechanical bond: the DNA rotaxane developed by the Famulok group of Universität Bonn.

### DNA as a building material

While DNA is primarily known as the information carrier of living organisms it has found recent application as a synthetic nano-scale engineering material. DNA has several key characteristics that makes the molecule highly suited for the creation of designed structures such as high stiffness, ease of modification and, most important the unmatched specificity and predictability of matched sequences as given by the Watson-Crick base pairing scheme.

The advantageous properties of DNA as an engineering material arise from the very same demands that the molecules role as the carrier of genetic information places on it. The molecule must be easily copyable which implies a high one to one base specificity (Watson and Crick 1953), it has to have the ability to be modified to carry various epigenetic signals (Slotkin and Martienssen 2007) (Robertson and Wolffe 2000)and yet the molecule has to be relatively chemically stable so as the backbone remain unbroken and the prime information carrying base pair sequence intact (Lindahl 1993). These demands suggest a highly adaptable and yet well ordered structure capable of higher order stable structure organisation.

#### *Base pairing and specificity*

The base-pairing rules first described by Watson and Crick (Watson and Crick 1953) in their revolutionary Nature paper has since the early seventies been exploited in the creation of novel DNA combinations from naturally occurring sequences (Cohen, et al. 1973). The basis of this versatile method of combining separate DNA fragments into one unit is the predictable nature with which two complementary single stranded DNA molecules combine into one unit bound together by the hydrogen bonds of the complementary bases. The complex from two hybridizing single strands is of the same B-DNA structure as the naturally observed double stranded molecule which allows precise prediction of the orientation of the adjoined molecules allowing large ordered structures as seen in Figure 1, (Qiu, Dewan and Seeman 1997). A major contributor to the stability of the double helix

structure is stacking between neighbouring bases. Indeed the favourable stacking characteristics of the GC pair are the main reason for the higher stability compared to that of the AT pair (Yakovchuk, Protozanova and Frank-Kamenetskii 2006). By non template *in vitro* synthesis of oligonucleotides custom hybridization regions can be created by widening the temperature and specificity range dictated by natural restriction enzymes.

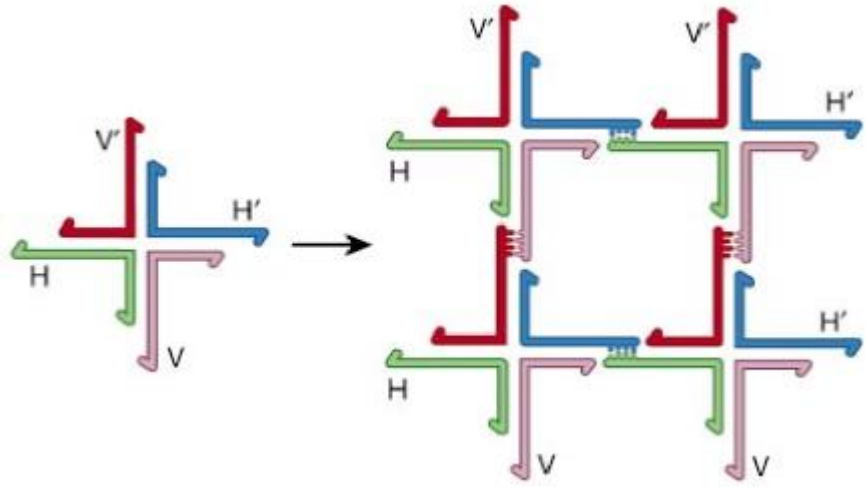



**Figure 1 Illustration of how four branch DNA nanotiles constructed from Holliday-junctions are able to self organize into larger periodic structures. The specificity and energy gain driving the reaction comes from the unique properties of the Watson-Crick base pair. Figure courtesy of N. C. Seeman,**

### *Stability*

While DNA is much more stable than its twin RNA owing to the absence of the latter's ribose 2' Hydroxyl group it is still very much affected by its surroundings. The major modes of damage to hydrolysed DNA are depurination of the sugar base, deamination of the base itself and general oxidative processes acting on several sites of the molecule. These decay processes are in most part effectively countered by a host of enzymatic repair processes in the living cell without which the genomic information would be rapidly corrupted at ambient temperatures. In the temperatures and conditions of normal DNA storage the molecule can nevertheless be expected to be stable with decay times due to non-enzymatic damage stretching over millennia. At higher temperatures and in the presence of DNA modifying enzymes the stability of the molecule might however be of concern over prolonged incubation times (Lindahl 1993). In order to increase resistance to enzymatic and/or chemical degradation synthetic nucleic acid analogs with similar base pairing properties to DNA may be employed. Examples include peptic nucleic acid, PNA, locked nucleic acid, LNA and 2' O-methyl modified RNA (Egholm, et al. 1993) (Kaur, et al. 2006) (Sproat, et al. 1989).

### *Functionalization*

Synthetic DNA is easily functionalized with various backbone or base analogs in the course of solid support synthesis and these introduced groups can be further modified in post synthesis treatments. The purpose of modification can be to attach ligands or enzymes to the nucleotide, to increase resistance to degradation or a number of other applications (Verma and Eckstein 1998).

### *Rigidity*

The ability to build wire frame like structures from DNA is dependent on the greatly enhanced stiffness of the DNA double helix compared to the single stranded molecule. The persistence length of double stranded DNA has been determined to 450-500 nm in common buffer conditions (Hagerman 1988). The stiffness can be further enhanced by the intertwining of two parallel helixes in a paranemic crossover motif seen in Figure 2. This structure can be envisioned as two parallel helixes where a crossover, also known as a Holliday junction, appears at each point the backbones meet (Shen, et al. 2004).

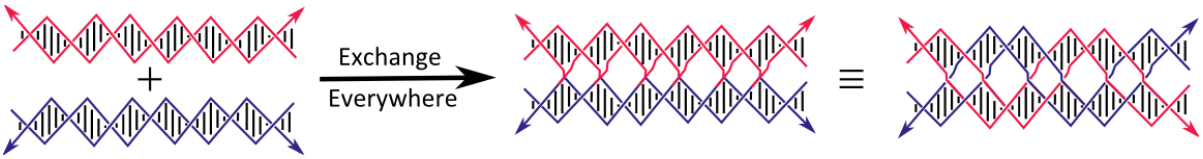



**Figure 2 A paranemic crossover motif of the PX type, the structure can be viewed as a pair of DNA helixes laying side by side and switching strands by a Holliday junction at each back bone intersection. This intertwining of two DNA double helixes greatly enhances the stiffness of the adjoined molecules (Shen, et al. 2004). Picture recreated from (Shen, et al. 2004).**

The tensegrity triangle is a rigid DNA motif made up from three interlocked DNA helixes whose high stiffness is derived from a combination of internal tension and stretching stresses. The three corners of the triangle are made up of four branch junctions, two branches forming the edges while the other two are open for interactions with the surroundings. The DNA helices are stacked in an up down manner at the corners so that each vertex is structurally identical to the two other. This flattening of the normally spread and floppy structure of the branched junction introduce an internal stress in the motif greatly increasing stiffness (Liu, et al. 2004).The tensegrity motif stands out as the first DNA structure with high enough rigidity to be able to self assemble into macroscopic crystals (Zheng, et al. 2009). A variant of this motif with edges made up of paranemic crossovers in place of single helixes has a further increased rigidity, although the principal advantage of this motif is the enhanced affinity of the double sticky ends presented at each interaction point (Zheng, et al. 2006).

Another approach to making DNA structures of higher rigidity are to employ DNA as a space filling structural material instead as a wire frame material. A major breakthrough in this area was achieved by Rothmund in 2006 when he described the directed folding of a long DNA strand into space filling motifs termed “DNA origami”. The technique works by guiding the folding of the long single stranded molecule into a pre defined pattern by the use of small staple strands forming helical double stranded domains with the longer strand (Rothemund 2006).

### *Branched DNA*

DNA is as it appears in nature is an essentially linear, if sometimes circular, molecule. If this was always the case the application of DNA as a structural material would be severely hampered by the lack of means to connect helices. What overcomes this limitation is the presence of motifs for DNA interaction that overcome this basic linearity.

The Holliday junction was first proposed in 1964 as part of a model for the yet to be fully characterized phenomena of meiotic conversion. The structure consist a heteroduplex where in

effect two adjacent double helixes are joined together, each substituting one native strand for the invading strand of the other. When the regions flanking the junction are symmetrical branch migration is allowed whereby the junction slides along the helixes, exchanging strands as it goes. When the two strands are different, say to contain different alleles of a gene, a transfer of information between the strands occurs (Stahl 1994).

Through the creation of branched junction with unsymmetrical arm sequences the migration of the complex can be hindered and the branch point locked in place across two linear molecules. The set of rules set forth by Seeman allows branched junctions with arms numbering between three and eight (N. C. Seeman 1982), and stable branched junctions with five and six arms has been demonstrated (Wang, et al. 1991) with a four arm junction shown in Figure 3.

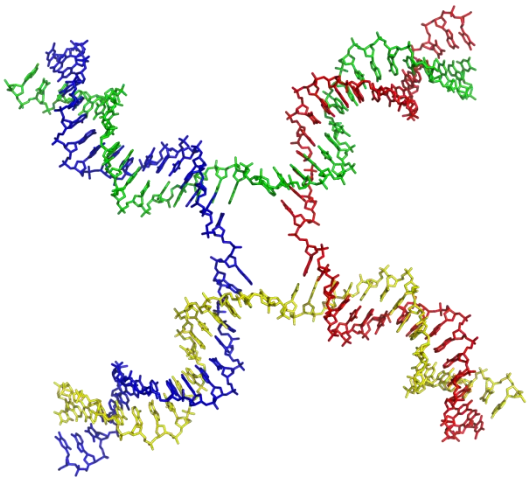


**Figure 3 The four arm branched junction consist of four strands and is one way to connect linear DNA into larger structures. The motif is a generalized Holliday junction lacking the symmetry that makes strand migration and chromosome pairing possible, locking the crossover in place (Stahl 1994). Picture from Wikimedia Commons, (Wheeler 2007).**

The generalization of the Holliday junction with crossovers at every interception point of two DNA helices yields the previously discussed paranemic crossover motif (See Rigidity). As the motif is paranemic, that is the composite helices may be separated without unwinding the complex molecule (Yagil 1991), the complex may be applied as a junction between two linear DNA helices. This method has the advantage that the joining regions can be made to arbitrary length and that closed loop strands can be joined (Shen, et al. 2004). A similar approach is to use specifically interacting RNA loop complexes hybridised to and replacing part of a DNA double helix. Two DNA structures equipped with complementary loops can then be bound together (Mayer, et al. 2008). Another method to join two linear helices is through the application of a strand invasion complex. Strand invasion of double

helical DNA is easily accomplished by complementary PNA *Peptide Nucleic Acid* oligos[1], which may be parallel or antiparallel to the DNA strand, and relies on the higher stability of the PNA·DNA duplex as compared to the native DNA duplex (Peffer, et al. 1993).

### *Curved DNA, AT tracts*

Apart from branching the ability to create a controlled bend in a structural material is an important feature as it allows an expanded range of structures to be made with fewer junction points. DNA has this ability to form intrinsic curvature from a structural repeat motif repeat termed AT-tracts. First discovered in *Leishmania tarentolae* kineto-plast DNA where the curvature manifest as unusually slow gel migration rate (Marini, et al. 1982), the bent DNA is due to periodic repeats of A·T pairs three to six nucleotides in length with a period of around 11 bp (Diekmann and Wang 1985). The AT tracts and the resulting circular curvature from the periodic repetition of this motif can be seen in Figure 4.

The bend of a six bp AT unconstrained repeat has in solution been decided to be 19° with the majority of the bend taking place at the borders of the AT repeat. This bending is mainly caused through stacking interactions with neighbouring base pairs resulting in a curvature of 4° at the 5' end, 10° at the 3' end and 5° spread evenly over the six base pair repeat (MacDonald, et al. 2001). A curve fitted to the calculated circularization efficiencies of self complementary AT tract segments of alternating five and six bp AT repeats peaks around 137 bp. This indicating a preferred curve of 55.2° per unit 21 bp repeat unit or a mean bend of 27.7° per an average repeat unit of 5.5 bp in circular DNA (Ulanovsky, et al. 1986), higher than the 19° bend per six bp repeat mentioned earlier.

[1] PNA or *Peptide Nucleic Acid* is a synthetic DNA analog utilizing a glycine peptide chain as a backbone as opposed to the sugar-phosphate backbone of natural nucleic acids. The molecule incorporates the standard DNA/RNA bases and forms Watson-Crick base pairs. This design offer several advantages over natural nucleic acids such as resistance to enzymatic degradation and the formation of stronger interaction complexes with itself and/or natural nucleic acids (Peffer, et al. 1993).

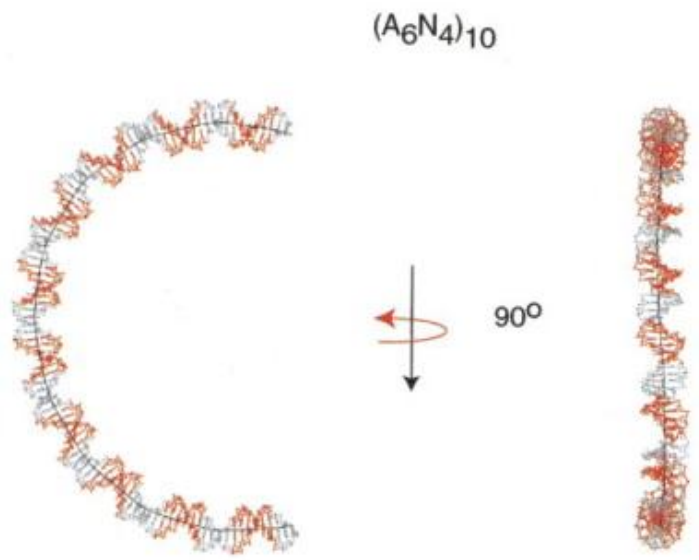



**Figure 4 Illustration of the curvature caused by ten repeated six bp AT tracts. Each AT tract contains a six base pair A·T repeat followed by four random pairs and causes a 19° bend of the DNA double helix. The bend of the AT-tract is shifted slightly out of plane and will cause the undulation of the backbone seen on the right. Illustration courtesy of Ponzy Lu.**

The curvature of the AT tract is stabilized by the presence of $Mg^{2+}$ ions (Diekmann and Wang 1985) and small DNA circles made up of periodic AT tracts lose their smooth curvature and assumes a kinked configuration when $Zn^{2+}$ is substituted (Han, et al. 1997) .

Alternate 5 and 6 bp AT repeats creates a 21 bp repeat motif that has zero helical pitch, that is, each 21 bp repeat contains two turns of the DNA double spiral and keeps the ends lined up in the same direction. From these segments a wide range of circular sizes can be created with circularization product down to 105 bp detected when 21 bp segments were self ligated (Ulanovsky, et al. 1986). The interruption of the 21 bp repeats patter with single gaps or functionalized parts are generally accepted with the flexibility of the DNA allowing bending of uncurved regions (Rasched, et al. 2007) (Synthesis). This feature allows the creation of interlocked and branched multi circle structures such as DNA rotaxanes (Ackermann, et al. 2010).

### *DNA nanorchitechtures*

To by DNA guide the assembly of other nano components, a concept termed DNA nanoarchitecture, was first proposed by in the early 1980:s as a mean to organize hard to crystallize proteins into regular three dimensional patterns for the purpose of x-ray crystallography, see Figure 5, (Seeman and Lukeman, 2005). The DNA architecture approach to crystallization where regular three dimensional DNA structures are used as a guide for the assembly of other particles, although simple in theory, has yet to be realised due to failure to produce the regular three dimensional DNA patterns themselves (Seeman and Lukeman, 2005). It is only recently that the first macroscopic DNA crystal from nano scale subunits have been produced employing the stiff tensegrity triangle DNA motif as a basis (Zheng, et al. 2009). This motif has already been utilized to order gold nano particles in a regular two dimensional DNA matrix (Zheng, et al. 2006) bringing hope of fulfilling the long stated goal of DNA nanoarchitecture.

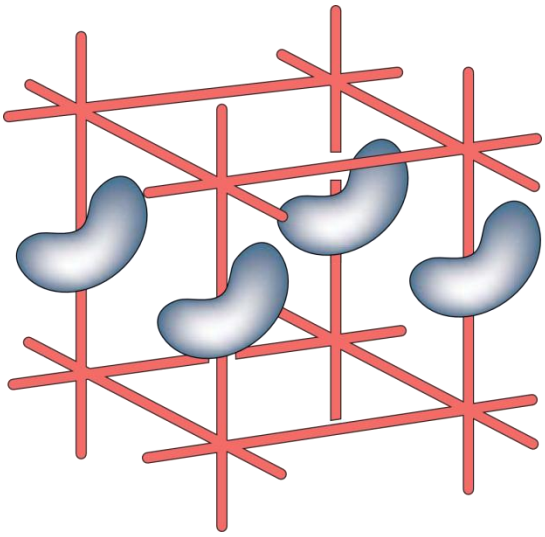


**Figure 5 DNA as a scaffold to guide the assembly and organization of other molecules was one of the first ideas to make use of the specificity of the DNA base pairing. It is however first recently that DNA motifs rigid enough to make the task possible has been developed. (Seeman and Lukeman, 2005). Illustration courtesy of N. C. Seeman.**

### *Kinetic self assembly*

Sequential assembly of DNA nanostructures can be programmed into the structural strands by a ligation chain reaction. This is accomplished by the sequential exposure of hybridization sites called "toe-holds" and the creation of the longest possible double strand (minimizing structure free energy) through the process of strand migration. A simple implementation of this idea are the molecular tweezers by Yurke *et al* in Figure 6. The process starts with a structure with exposed single strand regions open to hybridization. The single strand region is matched to another free structure capable of base pairing not only the single strand region but also to part of the already filled base pairs in the starting structure in such a way that the potential new double strand region is lower in energy, that is of greater length, than the old. By designing the sequences in such a way that the hybridization of the new strand and the simultaneous displacement of the old free a new toe-hold hybridization site the reaction can progress through an infinite number of self assembly steps, each energetically favoured over the previous. By this process structures of arbitrary size and complexity can be designed. Examples of created structures based on this design principle are bipedal walkers and dendritic networks (Yin, et al. 2008). Recently a programmable DNA walker has been produced which leverage of the toe hold principle (Gu, et al. 2010).

## DNA nanomachines

One of the first demonstrated DNA nanomechanical devices was a molecular tweezers shown in Figure 6 capable of repeated opening and closing upon an outside signal. The tweezers function by the controlled addition of short hybridizing DNA oligos to a partly single stranded structure forcing it to close up on itself. The hybridizing strands can be removed through the addition of antisense hybridizing strands returning the tweezers to its original open configuration. The fuel for the movement is thus provided by the free energy gain provided upon the formation of incrementally

longer hybridized structures (Yurke, et al. 2000). The disadvantage of using nucleotide oligos as fuel for a nanomechanical device is the accumulation of waste, hybridized strands, which cause declining efficiency as the reaction progresses (Liang, et al. 2008).

An improvement to this scheme that avoids the waste issue is to replace the hybridizing strand controlling the folding state of the DNA tweezers with a light driven reaction. This is accomplished by modifying the hybridizing strands with the light reactive molecule azobenzene. This molecule intercalates between the DNA bases in its cis-form but upon irradiation switch to trans-form disrupting the DNA double helix to which it is attached. The reaction is reversible and upon irradiation with a longer wavelength the molecule returns to its trans-form, again stabilising the DNA base pairing (Liang, et al. 2008).

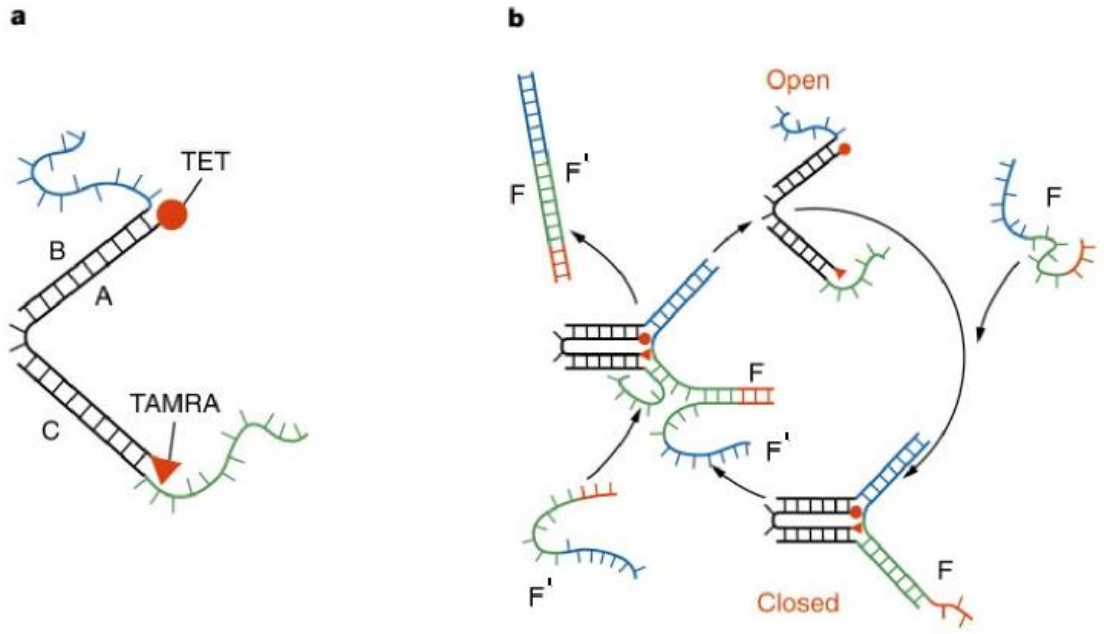



**Figure 6 The DNA molecular tweezers described by Yurke *et al.* a) The DNA molecular tweezers consist of: A, a DNA strand marked with a fluorophore (TET) and a quencher (TAMRA), B and C are two DNA strands partly hybridizing to the marked strand, forming loose single stranded ends at the duplex and a central nicked region acting as a hinge. b) The action of the molecular tweezers: The open molecular tweezers are closed by the addition of a shutting strand, F, which hybridizes to the free parts of strands B and C and forces the A strand into a bent configuration, bringing the fluorophore and quencher together. The F strand is not totally hybridized but maintain short single stranded region, a "toe-hold", this allows removal of the shutting strand by the addition of an antisense strand F' which replaces the B and C strands by strand migration, resetting the tweezers to the open state. Illustration courtesy of Andrew J Turberfield.**

A programmable path following DNA walker, Figure 7, has been constructed using the same basic design scheme as the previously discussed molecular tweezers. Here a walker made up of double stranded DNA with single stranded hybridizing regions termed hands is able to bind to a set of free hybridizing strands set in a DNA origami matrix. By the controlled addition of release oligos during

the run the walker releases and rehybridizes the path strands, moving from one handle to the other. On its course the walker passes a number of cargo pick up points with exposed toe-holds where a hand can grip and hold a cargo via a strand displacement process. The ability to pick up cargo, or the state of the pick up point, can be pre set through the addition of the hands of choice to the walker prior to the run. Through this approach a single walker type operating on a path with three pick up points can be programmed to produce eight distinct products (Gu, et al. 2010).

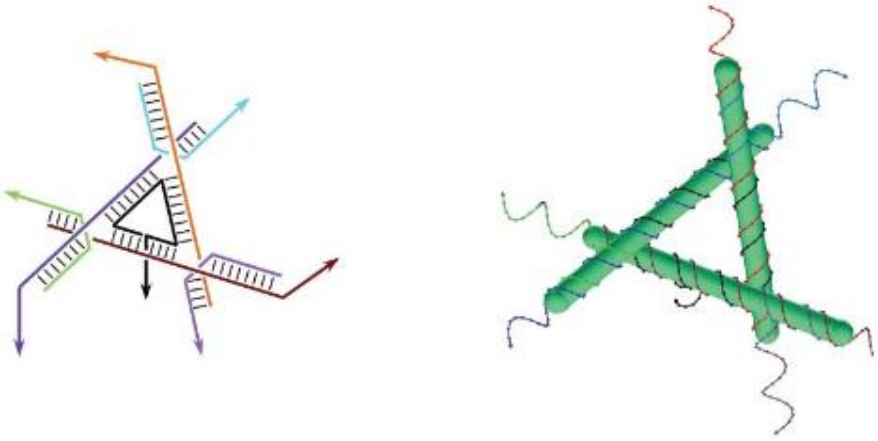


**Figure 7 A path following DNA walker capable of picking up cargo as programmed by selector strands, with line drawing showing constituent strands to the left and a three dimensional stick model right. The walker follows a pre set path by attachment and release of footholds complementary to the exposed single strands of the walker. By varying the attached strands a selection of cargo combinations can be formed. (Gu, et al. 2010). Illustration courtesy of N. C. Seeman.**

While the programmable DNA walker is able to follow and execute pre set commands the movement of the walker itself from position to the next still have to be directed by the addition of release oligos (Gu, et al. 2010) making a run time consuming and arduous. An autonomous path following one directional DNA walker has been designed by directed cleavage of the visited path oligos, preventing backtracking. The footholds of the path are set in a DNA origami matrix and a DNAzyme[2] functions as a foot that grips and subsequently cleaves the free oligo. Cleavage of the oligo strand is accomplished by the incorporation of a RNA base which the DNAzyme is able to recognize and hydrolyze. The cleavage of the foothold causes disattachment of the DNAzyme/oligo complex and the foot can peruse a new attachment point. By setting out a path of cleavable oligos a four legged walker can walk the path by sequential hybridization, cleavage and disassociation of the legs of the walker to the path oligos. A stop point for the walker is accomplished simply by substitution of the ribose by a deoxyribose in a group of footholds permitting attachment of the foot but not cleavage and dissociation (Lund, et al. 2010).

[2] A deoxyribozyme or *DNAzyme* is a catalytic DNA chain analogous to the better known ribozyme of RNA (Breaker and Joyce, A DNA enzyme that cleaves RNA 1994).

## DNA rotaxane

The Famulok group of the LIMES institute, Universität Bonn, has recently constructed a DNA rotaxane with DNA minicircles made up of AT tracts as a basic component (Ackermann, et al. 2010). A rotaxane is a mechanically interlocked structure where one or more macrocycles are threaded over a rod and restricted by bulky stoppers. As the molecules are bonded together by mechanical bonds, that is, the restriction is by sterical means, the parts are free to move relative to each other. (Bravo, et al. 1998). The DNA rotaxane demonstrated by the Famulok group and shown in Figure 8 is made up from DNA and is many times larger than the usual small molecule rotaxanes. In this system the macrocycle is a DNA minicircle and is threaded in place over a DNA rod by the matching of short single stranded regions of circle and rod creating a hybridized double stranded region. With the macrocycle caught the ends of the rod are blocked by the ligation of spherical stoppers made up of two minicircles connected by four arm junctions creating a wire frame structure. The addition of a release oligo substituting the macrocycle by duplex invasion releases the macrocycle from the rod. The spherical stoppers are large and stiff enough that upon release they form a steric barrier, keeping the macrocycle threaded over the rod and creating a stable rotaxane (Ackermann, et al. 2010).

**Figure 8 The DNA rotaxane by Ackerman *et al* consist of two parts, the central DNA minicircle and the DNA dumbbell which restricts the minicircle's movement to along its axis. The rotaxane is produced in steps by first threading of the minicircle and central DNA axis via complementary single stranded regions, secondly the wire frame stoppers constructed from two DNA minicircles connected by branched junctions are ligated to the ends of the central axis and lastly the minicircle is released from the central axis by the addition of release oligos. The stoppers are large and stiff enough to restrict the central minicircle to diffusion along the central axis thus forming a mechanical bond. Illustration courtesy of Damian Ackerman.**

The development of stable interlocked DNA nanostructures holds great promise in the development of DNA nanomachines. One might imagine a threaded minicircle modified with an enzyme keeping it in place while performing modifications of the base strand (Ackermann, et al. 2010). Another way to harness the freely moving parts of a rotaxane is to use the transcriptional movement of a polymerase acting on the components creating a nano engine. A promising candidate for this application is the T7 RNA polymerase, the action of which on DNA minicircles is the purpose of this thesis. A schematic representation of this design can be seen in Figure 11.

## RNAP T7

The RNAP T7 is a small single subunit phage RNA polymerase notable for its high promotor specificity, no need of external transcription factors and its very high processivity (Iwata, et al. 2000)

(Martin, Muller and Coleman 1988). Numerous studies of the T7 RNAP have been conducted resulting in deep knowledge of optimal promotor sequence (Breaker, Banerji and Joyce 1994) (Rong, He, et al. 1998)and proven strategies for immobilization and modification of the enzyme (Pomerantz, et al. 2005). This makes the RNAP T7 highly interesting as a component in synthetic nano scale machines (Pomerantz, et al. 2005).

### Promotor sequence

The T7 RNAP belong to a family of closely relate phage RNAP:s that share a common core promotor signature (Martin and Coleman 1987).The minimal functional T7 promotor sequence is **T AAT ACG ACT CAC TAT A** with truncations of, or mutations within, this motif drastically affecting the promotor affinity and or transcription rates. (Martin and Coleman 1987) Transcription initiates at the first G following this sequence with GGGAGA (Martin and Coleman 1987) or simply GGG (Diaz, Raskin and McAllister 1993) as the favoured initiation sequence.

Counted from the initiation site the T7 promotor is strictly conserved between base pairs -17 and + 6 of which the promotor binding domain is located from – 17 to -6 and the specificity domain, which selects between related polymerases is located from -5 to + 6 (Rong, He, et al. 1998). In a artificial evolution study of the T7 promotor some flexible positions are noted in the promotor binding region, the flexible bases are as follows -17 (T, A, C), -15 (A, T), -11 (G, C) (Breaker, Banerji and Joyce 1994).The binding domain is largely conserved in the related group of phage polymerases, including T7, T3 and SP6 among others, and is recognized as a double stranded DNA helix. Interactions with the specific bases take place on the non template strand bp -11 and -10 but shift over to the template strand from bp -9 to -6/-5 (Rong, He, et al. 1998).

The region proceeding the consensus promotor region (Martin and Coleman 1987) seem to have little impact on promotor function with no sequence preferences for the region -23 through -18 found (Breaker, Banerji and Joyce 1994). It is however suggested by an early study that the most efficient T7 RNAP promotors conform to the consensus sequence GAAAT at positions -22 through -18 (Dunn, Studier and Gottesman 1983).

### PTH terminator motif

The PTH terminator is a short T7 RNAP terminator first discovered in human prepro-parathyroid hormone (PTH) gene and is unusual in the sense that the termination is not caused by the formation of RNA secondary structure. Instead the resolution of the transcription complex is thought to be caused by a DNA duplex configuration. The terminator sequence is **HATCTGTTTT** (where H is A, C, or T) on the coding strand with termination occurring with preference for a site five nucleotides downstream from this motif. The termination efficiency is estimated at around 40 % with gaps, nicks or super coiling of the template having a strong negative impact on termination efficiency (He, et al. 1998). The PTH terminator is highly interesting for the application in DNA minicircles as the motif fits in the reverse strand random five base region between the Thymine runs. This enables the integration of a terminator in a ring shaped template without disrupting the helical bend of the AT tracts.

## Promotor independent transcription

Even though the T7 RNAP show extremely high specificity of its double stranded promotor (Martin and Coleman 1987) there are numerous examples of promotor independent transcription from single stranded oligo nucleotides (Unrau and Zaher 2004) (Rong, Durbin and McAllister, 1998) (Marras, et al. 2004) (Cazenave and Uhlenbeck 1994).

The T7 RNAP has the ability to with high affinity bind to single stranded DNA or RNA oligos and processes these with an efficiency approaching that of double stranded DNA with a promotor sequence (Unrau and Zaher 2004). The action of binding to the single stranded RNA is dependent on the formation of a structure that may mimic the initiation bubble occurring at the start of a normal promotor dependent transcription. This happen when the oligo nucleotide is able to fold back upon itself or other present strands, and form a short loop structure stabilized at the 3' end by base interactions, shown in Figure 9. The polymerase is then able to bind to the loop directly in elongation state to elongate from the folded back 3' end using the complementary strand as template. The binding can be initiated by as little as one matching base pair and is most efficient when the binding is unstable and transient (Unrau and Zaher 2004).

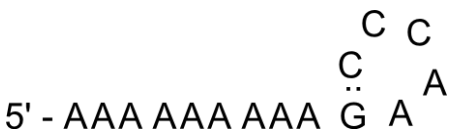



**Figure 9 Transient elongation initiation loop as described by Unrau 2004. The T7 RNAP bind to the loop formed by the switchback of the 3' end of the oligonucleotide stabilized by the transient C·G pair and subsequently elongates the oligonucleotide from the 3' end utilizing the oligo itself as template. The initiation of elongation by the T7 RNAP to these unstable loop complexes approaches that of a double stranded DNA promotor. Recreated from Unrau 2004, (Unrau and Zaher 2004).**

Supporting the notion that the polymerase binds in to a structure mimicking a transcription bubble is that the polymerase can bind in to synthetic transcription bubble termed a primed DNA duplex / RNA complex for initiation and efficient transcription, independent of the DNA sequence. The primed bubble mimics the active elongation complex and thus allows the RNAP to bind directly in elongation state. (Daube and von Hippel 1992).

Another issue that can produce unspecific elongation of a transcript is so called run around transcription. The T7 RNAP upon reaching the end of a template with a non template strand overhang can insert the free 3' end and thus utilize the old non template strand as the new template. The T7 RNAP may also traverse nicks, gaps and branch junctions in the template strand while producing a faithful complementary sequence to said strand (Rong, Durbin and McAllister, 1998).

The promotor independent elongation of oligo nucleotides can create problems through the formation of aberrant transcription products confounding the interpretation of transcription studies. Especially vulnerable to this form of unspecific interference are templates that have inefficient promotors or when high enzyme to template ratios are used (Cazenave and Uhlenbeck 1994). Another case when interference from unspecific elongation or product formation may be of concern is when oligonucleotides are employed as probes in real time monitoring of transcription. It has been found that the T7 RNAP may use molecular beacon probes as template, producing short

complementary RNA strands that trigger separation of fluorophore and quencher giving a false, target independent, florescence increase (Marras, et al. 2004).

The elongation of oligonucleotides can be hindered by the design of stable 3' end product complexes with mismatched last nucleotide. This has the simultaneous effect of closing the back loop mimicking a transcription bubble seen Figure 9 in while also denying a matched base pair from which to start elongation. (Nacheva and Berzal-Herranz 2003) The unspecific triggering of oligonucleotide probes may be hindered by the use of synthetic nucleic acid analogous not recognized by the enzyme such as 2' O-methyl RNA (Marras, et al. 2004).

### T7 RNAP as a molecular motor

The T7 RNA polymerase has several characteristics that make it interesting as a molecular motor in synthetic DNA nanostructures. Prominent are the high specificity of the RNAP T7 for its promotor sequence, a highly stable transcription complex, the ability to halt and restart the transcriptional movement by force application or withholding of NTP:s and a developed force of propagation matching or larger than that of natural rail following motor proteins such as kinesin.

Control over the transcriptional progression of a T7 RNAP has been achieved by the withholding and selective addition of NTP:s necessary to form a transcription product. Upon reaching a stretch of template demanding NTP species not present in the reaction mixture, the polymerase halts and can upon addition of the necessary monomers resume transcriptional movement. This allows tight and template dependent motor control of the polymerase (Pomerantz, et al. 2005). Mechanically controlled pausing of a transcription has been achieved for Escherichia coli RNA polymerase. When the force applied to a transcribing complex exceeds the force developed by the polymerase the enzyme reversibly halts its progression along the DNA strand. Upon removal of the opposing force the majority of the stalled complexes are able to resume transcription of the template (Yin, et al. 1995). A light switch controlling the initiation of transcription from the T7 promotor may be built by the inclusion of the small organic molecule azobenzene into the promotor motif (Liu, Asanuma and Komiyama 2006). This molecule intercalates between the DNA bases in its trans-form but upon irradiation switch to cis-form disrupting the DNA double helix to which it is attached. The reaction is reversible and upon irradiation with a longer wavelength the molecule returns to its trans-form, again stabilising the DNA base pairing (Liang, et al. 2008).

Accurate measurements of linear force or torque of a transcribing RNAP T7 has not been performed but values from studies on the related Escherichia coli RNA polymerase can be expected to be in the same range (Pomerantz, et al. 2005). The force developed by a transcribing Escherichia coli RNA polymerase has been measured by monitoring the maximum force tolerated by an active transcription complex without causing stalling of the complex. The force of transcription was found to be in mean 14 pN, significantly larger than the force measured for the cytoskeleton motors myosin and kinesin and larger than the loads calculated opposing transcription in vivo. The measurement was made with an optical trapping interferometer, with which a calibrated force opposing that of transcription was applied to an attached styrene bead (Yin, et al. 1995). The torque developed by the Escherichia coli RNA polymerase while transcribing from a double stranded DNA helix has been estimated to >5 pN nm. This value was obtained by comparing the transcription velocity of surface

immobilized polymerases acting on dsDNA with a florescent bead attached to that of the known transcription velocity of the unrestricted polymerase. The difference was assumed to be due to viscous drag on the attached bead from which the torque was calculated (Harada, et al. 2001).

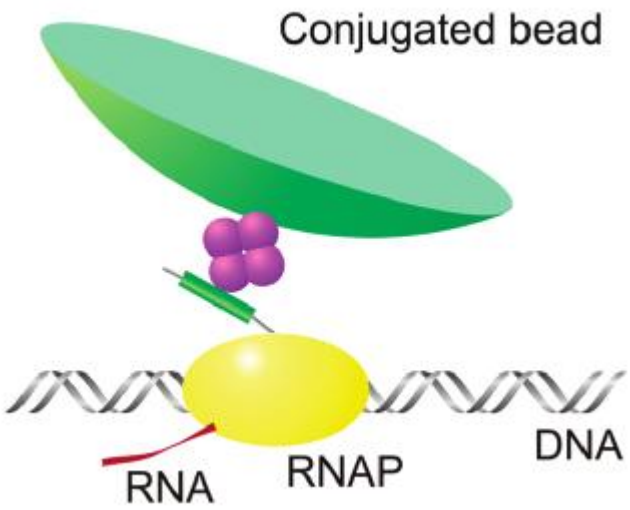



**Figure 10 The T7 RNAP as a rail following molecular motor on double stranded DNA analogous to myosin-actin systems found in cells. The polymerase pulls the conjugated along as it transcribes from the DNA template, the energy for the action being provided by the polymerization of RNA (Pomerantz, et al. 2005). RNA polymerases acting in this configuration has been recorded to exert in excess of 14 pN, significantly more than that provided by the kinesin and myosin systems (Yin, et al. 1995). Illustration courtesy of William McAllister.**

The motor properties of the T7 RNA polymerase have been applied to produce a linear driving force. The application of the enzyme as a rail following molecular motor has been demonstrated by monitoring the progression of nano-dot coupled RNA T7 polymerases on a combed DNA template (Pomerantz, et al. 2005) with the enzyme, template and cargo bead seen in Figure 10. This application echo the natural rail carrying transport systems, such as of the kinesin-microtubules system, responsible for tasks such as muscle contraction and organelle transport. Kinesin walkers have been demonstrated to be able to cooperatively transport large cargos in uniform fashion over a surface by simultaneously attach to several isopolar immobilized microtubule (Böhm, et al. 2001). The adaptation of the T7 RNAP for this task promise greatly simplified substrate production and integration with DNA nanotechnology (Pomerantz, et al. 2005).

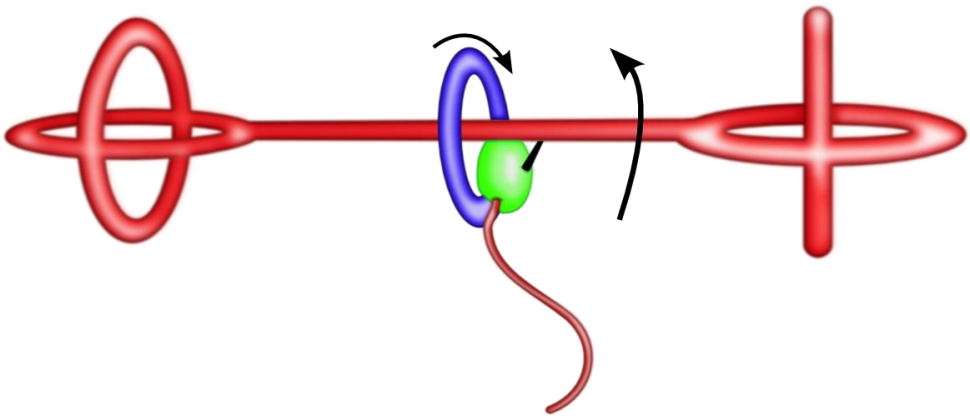


**Figure 11 Schematic design of a hypothetical rotary DNA-nanomotor analogous to the linear motor in Figure 10 created by Pomerantz *et al* (Pomerantz, et al. 2005). The motor consist of a DNA minicircle (blue) threaded over a DNA dumbbell (red) creating a rotaxane. Attached to the dumbbell is an RNAP T7 (green) whose transcription of the minicircle causes internal rotation of the system. The energy driving the reaction comes from the polymerization of NTP:s into RNA (brown) which in this system is purely regarded as waste and could be removed by the addition of RNases.**

Another area where RNA polymerases may challenge natural motor complexes is in providing rotational force. Here work has been done with the $F_1$ fragment of the adenosine triphosphate (ATP) synthase to drive a synthetic nano propeller made up of Ni rods with continued rotation observed for over two hours (Soong, et al. 2000). The ATP synthase $F_1$ has also been given an off/on switch by incorporating a zinc binding domain, providing ATP dependent rotation inhibited in presence of zinc (Liu, et al. 2002). The ability of the polymerase to provide torque on a linear substrate (Pomerantz, et al. 2005) (Harada, et al. 2001); to translate linear force into circular motion by operating on a circular substrate (the purpose of this work) ; the ability to control transcription initiation by light (Liu, Asanuma and Komiyama 2006) and transcription progression via template sequence and substrate levels (Pomerantz, et al. 2005) or mechanical means (Yin, et al. 1995) makes the RNAP T7 an attractive option to provide nano scale torque. A hypothetical rotary DNA nanomotor can is presented in Figure 11 and consist of a DNA dumbbell with a threaded DNA minicircle on which an RNA polymerase acts. The free rotational properties of the created DNA rotaxane coupled with the transcriptional force developed by the active polymerase causes rotation of the systems parts relative to each other.

## Molecular beacons for Real-Time monitoring of transcription

Molecular beacons are a form of dual labelled oligonucleotide fluorescent probes (Mackay and Landt 2007)where the free probe assumes a hairpin configuration, keeping the attached fluorophore and quencher into close proximity, Figure 12. Upon exposure to the single stranded target a linear double stranded complex forms separating the fluorophore from the quencher and providing a fluorescent read out (Tyagi and Kramer, 1996). The basic molecular beacon probe design has with some modifications been utilized to monitor real time transcription (Marras, et al. 2004) and rolling circle replication events (Nilsson, et al. 2002).

While the main rationale for the stem loop configuration of a molecular beacon probe is to provide stable quenching of the fluorophore (Tyagi and Kramer, 1996), the hairpin configuration of the free probe has the simultaneous advantage of increased specificity to its target compared to a linear probe design. This as the stem provides an energy barrier that the hybridisation of the probe loop to the target must overcome. By fine tuning the relation of stem to loop length of the hairpin, molecular beacons capable of discrimination between two target strands differentiated by as little as one nucleotide can be achieved at a given temperature. The molecular beacons can be labelled with a range of dyes allowing multiplexing analysis of several targets in one reaction (Tyagi, Bratu and Kramer, 1998). This combination of low background, multiplexing ability, and high target selectivity makes molecular beacons an ideal tool which to monitor real time transcription events (Liu and Feldman 2002).

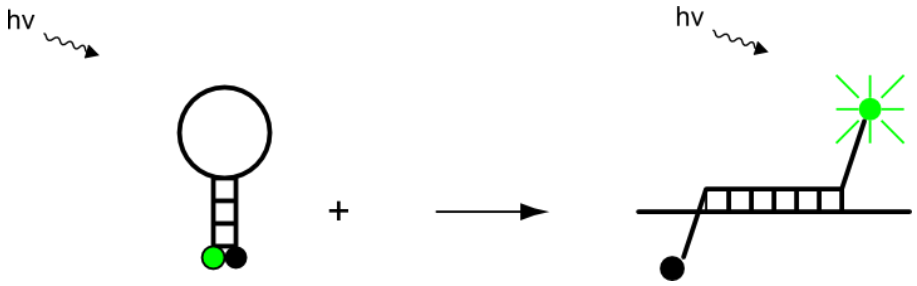



**Figure 12 The principle of a molecular beacon fluorescent probe: The unbound probe exist in a stem loop configuration keeping the quencher and the fluorophore in close proximity. When the probe is exposed to its target the loop sequence hybridizes to the target, forcing the quencher away from the fluorophore, allowing fluorescence upon irradiation (Tyagi and Kramer, 1996). Picture recreated from (Marras, Tyagi and Kramer 2006).**

A DNA molecular beacon is usually 15 to 30 nucleotides in length with GC rich stem around seven base pairs in length. The stem length and sequence should be chosen so that the stem loop configuration is stable at the intended detection temperature; in general a more stable stem provides higher specificity at the expense of slower hybridisation characteristics. The hybridizing loop sequence should have a melting point of seven to ten degrees Celsius above the detection temperature if single nucleotide discrimination is necessary. More stable probes which accept some target variability and give a higher fluorescence increase upon target exposure are obtained by raising the melting point of the target probe hybrid. The length of the loop is thus decided by the intended use of the probe as well as the target sequence. The structure of the designed probe can be checked with a DNA folding program in order to ensure that the intended secondary structure is produced. It is important that the probe is in its intended hairpin configuration so that the fluorophore is held close to the quencher; in contrast short two to three nucleotides long stems in the probe loop region have no adverse effect on probe performance (Vet and Marras 2005).

Molecular beacon florescent probes have successfully been applied in the real time monitoring of RNA transcription events (Liu and Feldman 2002) (Marras, et al. 2004). Problems have however been observed when using standard DNA molecular beacons due to the RNA polymerase using the probe as a template to produce short complementary RNA:s opening up the hairpin structure and increasing fluorescence. This issue is easily avoidable by utilizing the RNA analogue 2’ O methyl RNA as a backbone which is not recognized by the polymerase (Marras, et al. 2004). As the RNA·2’ O-methyl RNA duplex is more stable than the corresponding RNA·DNA or RNA·RNA duplex (Tsourkas,

Behlke and Bao 2002) the probe stem and loop lengths should be shortened compared to that of a standard DNA molecular beacon to obtain similar hybridization kinetics (Marras, et al. 2004) (Nilsson, et al. 2002).

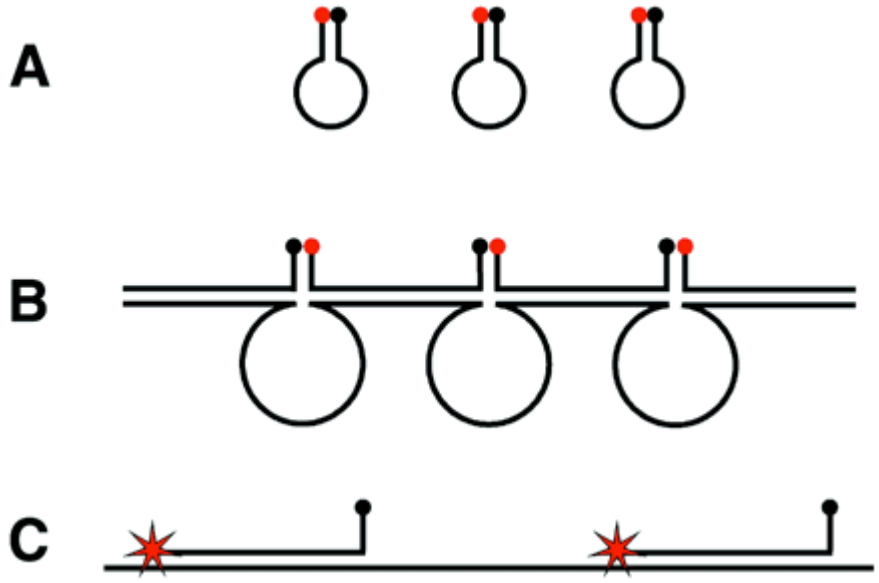



**Figure 13 Rationale for a stem hybridizing molecular beacon. When an ordinary molecular beacon (A) is used against a repeated target such as rolling circle products, inter molecular quenching may be problematic. This as the hybridization to the target opens the stem of the probe, allowing hybridization to adjacent probe stems on the same target strand, the loose single stranded target scrunching up between the stiffer target/probe double helix regions (B). A solution to this problem is to make one side of the stem sequence target complementary, thus avoiding inter molecular stem interactions (C). The resulting stem hybridizing molecular beacon has identical native configuration as the regular probe (A (Nilsson, et al. 2002) . Illustration courtesy of Mats Nilsson.**

When the target sequence of the molecular beacon is repeated in the target strand such as is the case when monitoring rolling circle transcription products problems with inter molecular quenching might arise, Figure 13 B. This is due to the loose stem sequences left when the loop part of a traditional molecular beacon hybridizes with its target strand. As the beacons lie lined up on the target strand separated by loose single stranded DNA, the loop strands are able to come into contact and form an inter beacon stem while still attached to the target. This brings together the fluorophore and quencher, greatly diminishing the fluorescent gain from probe target hybridization. The inter probe quenching can be avoided by a modified beacon design where also the stem part of the probe is designed to hybridize to the target strand, Figure 13 C. It has been found that making the quencher side of the probe target-complementary gives a higher fluorescence gain not only compared to the alternate fluorophore hybridization configuration but also compared to the normal free stem. This may be due to lower interference with the fluorophore when the quencher is held close to the duplex while the fluorophore is instead quenched by base interference if held close to the duplex (Nilsson, et al. 2002).

# Design

## Circle Family AK4

In order to be able to study the size effect on transcription from DNA minicircles a family of circles sharing a common promotor motif was designed ranging from 105 to 210 bp in increments of 21 bp. The insertion of a terminator motif give rise to a range of size dependent and size independent transcription products where the size independent products are identical for all templates featured in this study, allowing identification of transcription products on gel. The general layout of the circle family is given in Figure 15.

The minicircles are made up of 21 bp repeat units containing two AT tracts, six and five base pair in length, interspersed with five random bases giving an overall bend of around 40 degrees (MacDonald, et al. 2001) and a neutral helical pitch (Ulanovsky, et al. 1986). In order to increase specificity and stability of the unligated circle the five random bases interrupting the AT tract were chosen to have a high melting temperature and to be unable to form dimers with other distance elements. All 256 AAAAA SNNNS AAAAA(S = any base but A, N = any base) combinations were evaluated in regard to melting temperature using SantaLucia nearest-neighbour thermodynamics (SantaLucia 1998). The remaining combinations were checked for dimer formation utilizing a NUC44 scoring matrix leaving 64 SNNNS combinations with high melting temperature relative to the starting set and with low dimer formation with the set and anti set. These sequences can be found in Appendix 1: Distance elements.

The circle family was designed with two spacer segments of two 21 bp repeat units each, one variable length segment allowing shortening by a 21 bp repeat, and one constant length segment, Figure 14. By varying the hybridizing sticky end regions the segments can be combined to create the circles in Figure 14 consisting ten to five 21 bp repeat units spanning in circumference between 210 and 105 bp. A blunt end linear reference fragment was designed by omission of one spacer segment and filling out of the adjoining segments sticky ends. The circle family is thus built from one promotor containing segment, one terminator segment and two spacer segments. All segments and oligos making up the templates can be found in Appendix 2: AK4 circle family.

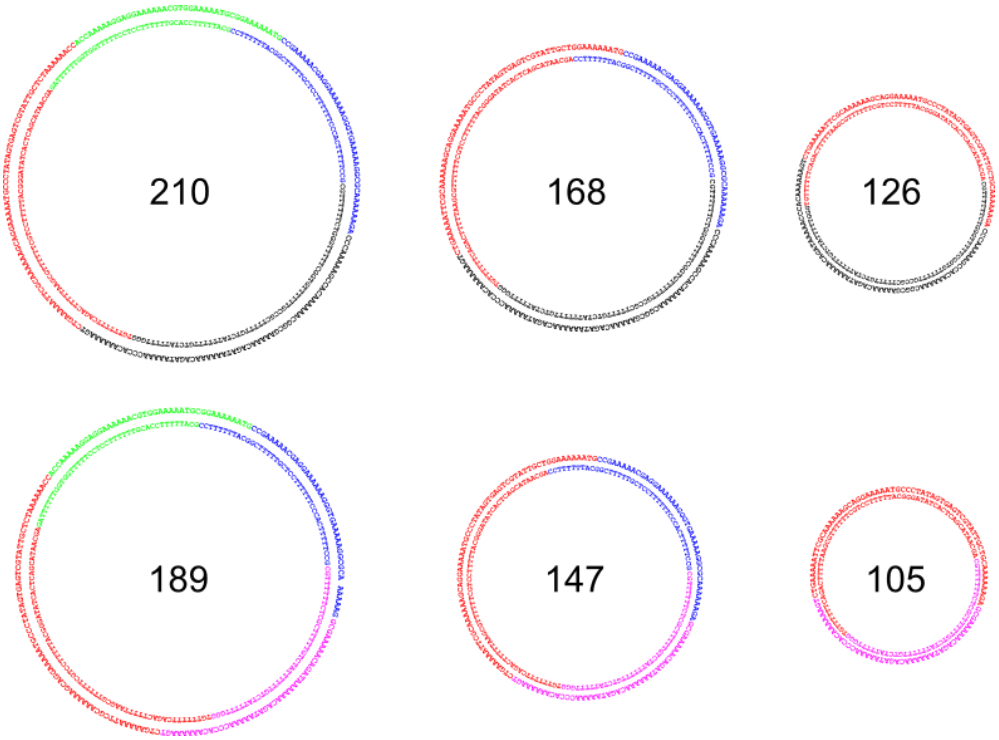



**Figure 14 The AK4 circle series with relative sizes to scale. The series of six circles spanning from 210 to 105 bp in increments of 21 bp is constructed from four unique segments, one of which has variable length (red, green, blue, black/pink). The circles decrease 48 bp per step from left to right, that is looses one omittable segment, and 21 bp from up to down by shortening of the variable length segment. The order of the segments is decided by which type of sticky end connection is chosen.**

The promotor region contains the T7 core promotor **T AAT ACG ACT CAC TAT A** (Martin and Coleman 1987) flanked by a downstream initiator and an upstream padding region and substitute one 21 bp repeat unit. The initiator region was chosen as **GGGAGA** (Martin and Coleman 1987) with the upstream region truncated from GAAAT (Dunn, Studier and Gottesman 1983) to **AAT** in order to fit the promotor region into one repeat region. The upstream region has been shown to have little importance to promotor efficiency (Breaker, Banerji and Joyce 1994) and here mainly serve as distance element buffering the start of the core promotor from the sharp bend in the first two base pairs following the AT tract (MacDonald, et al. 2001). The complete sequence of the promotor region, with the core promotor in bold, is thus: AAT **TAA TAC GAC TCA CTA TA**G GGA GA.

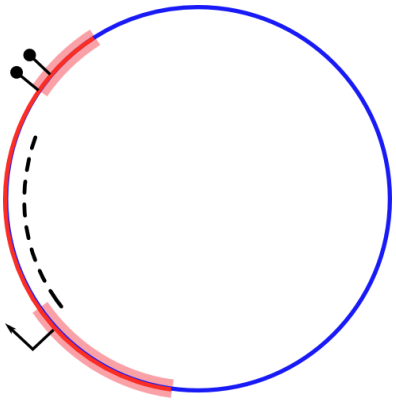


**Figure 15 Illustration of the design of the AK4 family of circles and circle fragments. The templates are made up of two principal parts, a constant region marked in red and a variable region in blue. The sizes of the circles are decided by the length of the variable region, which can be gapped forming a linear fragment, while the constant region is common to all templates. The constant region contains the promotor, → (arrow), and, the terminators, • (bullet), with the motifs shaded in red. Situated between the promotor and terminator is the region which when transcribed forms the target for the molecular beacon mbAK4P2, - - - (dashed) .**

Two PTH terminator motifs are integrated in tandem on the reverse strand of the terminator sequence giving a theoretical overall termination efficiency of 64 percent (He, et al. 1998). The distance between the initiator and first terminator falloff point is 56 bp with an additional 12 bp to the second terminator falloff point. This forms a constant termination signature of two circle size independent termination products of 56 and 68 nucleotides in length at 40% and 24% of the total product amount. These size constant products are followed by circle size dependent run around products at increasing length and decreasing molar ratios. The blunt end fragment produces a run off product of 146 nt in length.

## Molecular beacon mbAK4P2

The molecular beacon mbAK4P2 is directed against the P2 promotor segment of the AK4 circle family. Specifically the probe is directed against an area of the circle stretching from the promotor initiation site to the first ligation site of the promotor segment (nucleotide +1 to +25 of the transcribed product), Figure 16. This to ensure that all promotor clearing transcripts are detected as unspecific falloff is observed at ligation sites (own observation). The beacon is made from 2' O-methyl RNA to prevent the T7 RNA polymerase to use it in promotor independent transcription (Marras, et al. 2004) and the 3' quencher carrying part of the stem is designed to hybridize to the target in order to prevent inter molecular quenching (Nilsson, et al. 2002).


**Figure 16 Schematic view of the 2' O-methyl RNA molecular beacon mbAK4P2: The probe targets a transcribed target (red) from +10 to +25 bp from transcription initiation, fitting into a short stretch of the template (***) from the initiator sequence (i) to the first ligation point (gap). The molecular beacon is of the stem hybridizing type in order to prevent inter molecular quenching, with the target complementary loop and 5' stem shown in blue.**

Because of the special requirements in placing of the hybridization site, the use of 2' O-methyl RNA, and the hybridizing stem, the use of standard molecular beacon design software was deemed unsatisfactory. The probe was manually designed and different probe designs were evaluated in regard to the inter-molecular hairpin formation of the probe as well as the hybridization to the target strand. The evaluations were carried out with the Vienna server (Gruber, et al. 2008) software RNAfold (Hofacker, et al. 1994) and RNAcofold (Bernhart, et al. 2006) for intra and inter molecular folding patterns respectively and the typical folding of the probe in its native configuration is seen in Figure 17. As the 2' O-methyl RNA has folding properties closer to RNA than DNA (Sproat, et al. 1989) Andronescu model RNA parameters (Andronescu, et al. 2007) were used in the evaluation, the temperature was set to 37 °C. The free energies for hairpin formation and target hybridization as well as probe secondary structure was considered and compared to that of known (Marras, et al. 2004) (Nilsson, et al. 2002) (Tsourkas, Behlke and Bao 2002) functional 2' O-methyl molecular beacons. As the 2'O methyl RNA produces stronger hybridisation complexes with itself and RNA compared to standard RNA (Tsourkas, Behlke and Bao 2002) the structure was designed to be only marginally stable as indicated by RNA folding analysis with the final designs intended hairpin configuration frequency in the thermodynamic ensemble at 22.62 % with a free energy of - 2.10 kcal/mol upon folding. The energy gain upon hybridization with the target sequence was calculated to -14.52 kcal/mol and a melting temperature of 30 °C. Finally the dot plots of probe, probe plus target strand, and target strand with excised hybridization region was studied to ensure intended secondary structure and rule out the existence of secondary hybridization sites in the target strand. The fluorophore was chosen as fluorophore FAM (6-carboxyfluorescein) attached at the 5' end with DABCYL as matched quencher as given by several sources (Nilsson, et al. 2002) (Liu and Feldman 2002) (Marras, et al. 2004).

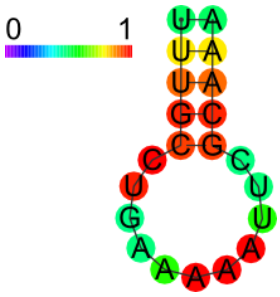



**Figure 17 The intra molecular folding pattern for the molecular beacon mbAK4P2 as calculated with RNAcofold (Bernhart, et al. 2006), Andronescu model RNA parameters (Andronescu, et al. 2007) at 37 °C. The probability of the given configuration at each site is colour coded from 0 (lilac) to 1 (red). The configuration shown has the dominant frequency of 22.62 % in the thermodynamic ensemble with high probability positions in red and moderately probable positions in green.**

# Result and discussion

## Synthesis

The synthesis of the DNA templates was carried out as described by Ackerman *et al* (Ackermann, et al. 2010) with the purified circular products shown in Figure 19.The efficiency of circularization was calculated by comparing the area under peak for the analytical HPLC report of each template corrected for molar absorption ratios, Figure 18. The results show a successful synthesis of circular templates down to 126 bp in length, albeit at a low efficiency for the smaller circle sizes. The decrease in synthesis for the circular templates as compared to the linear control is expected as the circularisation of the DNA helix lowers the molecules entropy (Ulanovsky, et al. 1986). The effect is less pronounced at the larger unstrained circles of 210 to 180 bp while the synthesis efficiency of the 126 bp circle is an order of magnitude lower than for the linear control. To a large extent this decrease can be explained by the formation of circular double fragment products where two linear fragments with matching sticky ends combine to form a relaxed product of double the size rather than the intended strained single product. Somewhat surprising is the low yield for the 168 bp circular template as this size has been shown to be the favoured configuration (Ulanovsky, et al. 1986). A possible explanation for this deviation is the insertion of the 21 bp straight promotor segment which might skew the optimal size upwards due to over bending of the remaining segments upon circularization.

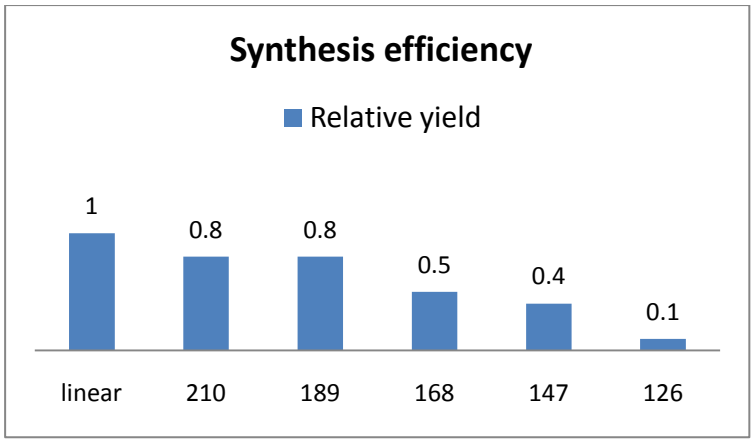



**Figure 18 Molar synthesis efficiency from 200 µl ligation reaction o the circle series from 210 to 126 bp relative to the 178 bp linear control as estimated by analytical HPLC. The molar yield is relatively unaffected by the circularisation for the larger 210 and 189 bp circles after which a sharp drop down to ten percent relative efficiency is noted for the 126 bp circle. Somewhat surprising is the low efficiency for the 168 bp circle which has been shown to be the favoured size when circle segments are self-ligated (Ulanovsky, et al. 1986). This may be due to the insertion of the unbent 21 bp promotor region increasing the unstrained size upwards. The sharp drop in synthesis efficiency noted for the 126 bp circle is caused by an abundant formation of 252 bp double circle product formed by self ligation of two linear 126 bp segments.**

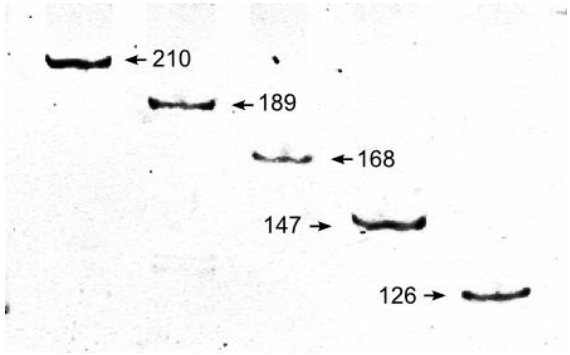



**Figure 19 PAGE Gel of purified circles illustrating the stepwise decrease in size. The whole family of circles is made up from five segments sharing four matched sticky end overhangs. The size of the circle product is varied by switching between different sticky end overhangs for the circle constant region allowing different ring closing reactions.**

The series of circles presented in figure two is the result of the successful implementation of the circle family AK4 modular circle design. The modularity of the design allows synthesis of a wide range of circles from a small number of precursor synthetic oligos. This translate into cost reductions as fewer oligos has to be stocked and ordered and better ability to discern differences between circles as being an effect of structure rather than material from different batches. As seen in Figure 19 the synthesis of small circles is very difficult with yields an order of magnitude smaller than for the linear control. This difficulty to readily produce the large template amounts needed necessitated the exclusion of the 126 bp circle from the polymerase kinetic study performed later in this work. The synthesis of the 105 bp circle was not attempted due to earlier in lab experience of low yields for very small modified circles.

## Transcription study

A transcription study of the circle family was carried out on the circle family ranging from 210 to 126 bp in circumference with a linear template as control. The transcription products were visualized on

a denaturising polyacrylamide gel (10% PAGE, 8M urea) and can be seen in Figure 20. Clearly seen in all the transcription reactions is the double terminator motif indicating the promotor/terminator cassette shared by all the templates giving a common reference point. The termination products visualized on the gel follow the expected length of 68 and 56 base pairs with molar intensities of 24 and 40 percent of the total product amount. Above the double terminator motif there are clearly discernable transcription products of length greater than the circle circumference for the 210 to 168 bp circular templates. For the smaller 147 and 126 bp templates there are indications of bands much longer than the template circumference but these are much longer than the first falloff length and very diffuse. There is no indication of RNase activity or unspecific transcript elongation as seen by the crisp appearance of the double terminator bands and the virtual absence of transcription products longer than the falloff product for the linear template.

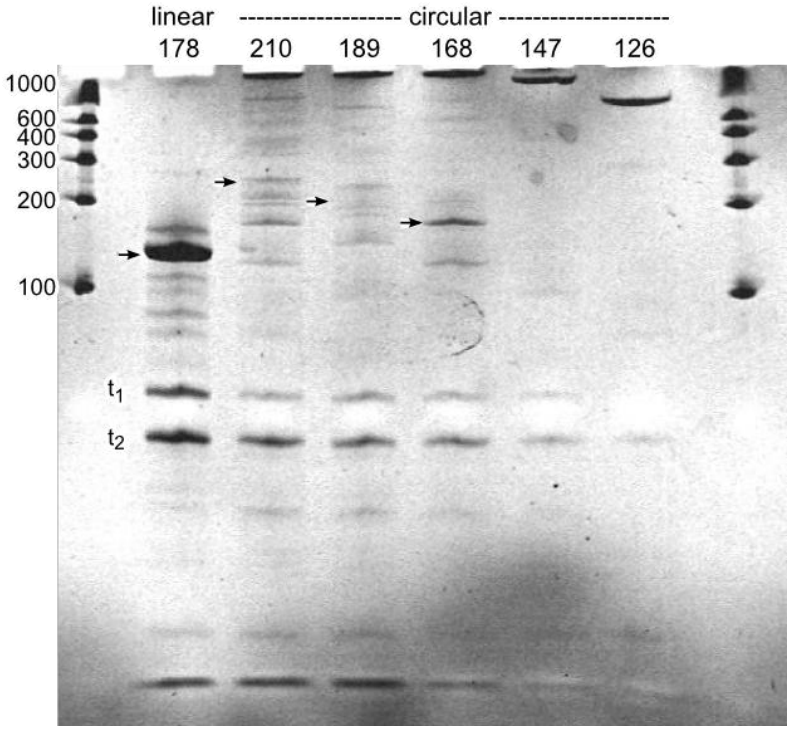



**Figure 20 Denaturating polyacrylamide gel of transcription products from circular templates from 210 to 126 bp and a 178 bp linear control, for size comparison a RNA ladder from 1000 to 100 nt is included. Marked, $t_1$ and $t_1$, is the double terminator signature from the tandem PTH terminator motif with expected lengths of 56 and 68 nt in length. The $t_2$ band is clearly of lesser intensity than the $t_1$ band confirming the 40/24 intensity ratio expected. Visible below the double terminator signature are shorter transcripts possible from aborted transcripts or other unspecific falloff prior to terminators. The size constant region below and including the double terminator motif confirms the identical promotor /terminator cassette present in all DNA templates and the decreasing intensity of the terminator bands gives a rough estimate of the initiation efficiency. Denoted by arrows (→) are the last visible transcription product of length at or shorter than the template circumference, for the linear template this corresponds to the 147 bp run-off transcript. For circle templates of 210 to 189 bp transcription products longer than this are clearly visible indicating rolling circle transcription. For templates of 147 and 126 bp in circumference the amount of product is to low to reliable identify transcription products above the double terminator motif. The sharp appearance of the terminator motifs and apparent absence of termination products longer than the fall of transcript for the linear template rule out unspecific transcription elongation as being the cause of the apparent rolling circle products.**

The identification of transcription products of greater length than the circumference of the circular templates in Figure 20 is conclusive evidence of rolling circle transcription with templates down to 168 bp. The common promotor and terminator motif is seen for all templates and the crispness of the terminator bands rule out unspecific elongation as the cause of the long transcription products. The expected ladder distribution is not seen, probably due to unspecific falloff outside of the terminators. This unspecific termination is best seen with the linear template where four major transcription products can be observed between the second terminator band and the run-off product. A decrease of transcription activity with template size is obvious with a large drop in transcription activity occurring when the template is circularized. It is however hard quantify the transcription intensities for the different templates due to the high background in the gel picture. Also the result is confounded by the long time of the transcription reaction where the transcriptional activity of the templates will be affected by factors such as enzyme exhaustion and substrate depletion. For more exact quantification and comparison of the transcription from the different templates a probe based real time assay is carried out.

## Real time monitoring of transcription from circular templates.

In order to more closely study the impact of the size of a circular template on the transcription efficiency a real time transcription study was carried out with the molecular beacon mbAK4P2 fluorescent probe as the indicator of produced product. The result of the transcription study can be seen in Appendix 3: Point kinetics with selected series presented in Figure 21 and resulting slopes from iteratively fitted linear regression in Table 1. As given by Table 1 the difference in slope for the larger templates appear to be very small with a large drop in transcription rate first observed for the 126 bp template. This observation is confirmed by a one way ANOVA of the variances between different templates versus the variance between replicates of the same template. In Table 2 it is shown that the in group SS (sum of squares) is in the same range as the between group SS for templates 210 to 147. Including the 126 bp template in the analysis roughly doubles the between group variability which translates to a rise in p-value above the 5% limit needed to reject the null hypothesis of all templates having the same transcription kinetics with above 95% certainty. The failure to properly discern the clear difference in transcription kinetics as observed by the larger product amounts produced for the bigger templates seen in Figure 19 with the real time monitoring is most likely due to the employed template concentration of 100 nM being near saturating conditions for the enzyme. To properly discriminate the efficiency of the RNAP T7 acting on the various circular templates a full kinetic study over a range of template concentrations is needed.

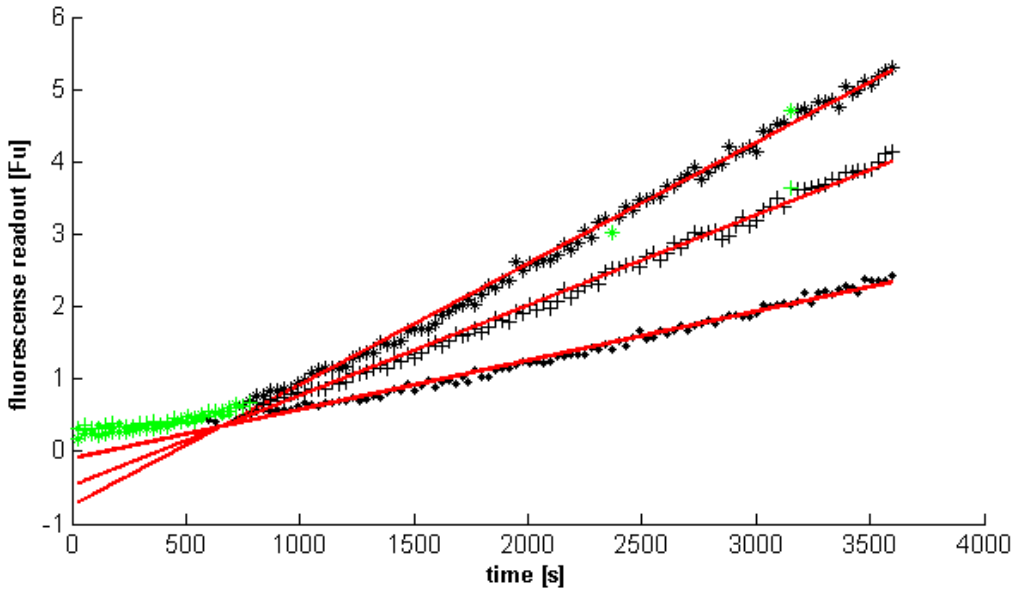



**Figure 21 Linear regression (fitted red line) of transcription kinetics single experiments on 4 pmol of template (210 *, 168 + and 126 • bp) in a volume of 50 µl illustrating a decrease in product formation rate with circle size. Outliers with respect to the linear regression line are iteratively observed and omitted as being outside the 95% prediction interval as given by nonsimultaneous observation bounds. Outliers are removed until the change in linear slope between two following iterations is less than 0.1 % . The 210 and 168 bp circle experiments have here been chosen to illustrate the general trend towards increased transcription rate with size, as seen in Table 1 the mean difference in slope between templates 210 to 147 is very small.**

**Table 1 Slopes for linear regression of transcription kinetics for 4 pmol of circular templates in a volume of 50 µl with 95% confidence interval and R-square goodness of fit measure. Duplicates for each template were performed and mean and variance calculated. As can be seen by the ANOVA in Table 2 the three largest cirlce sizes produces indistinguishable transcriptions kinetics at the sole monitored template concentration.**

| Size | Replicate | Slope | conf. Interval | | 95% R-square | Mean | Variance |
|---|---|---|---|---|---|---|---|
| bp | | | mRFU/s | | | | |
| 210 | 1 | **1.7** | 1.7 | 1.7 | 1.00 | 1.8 | 1.5E-02 |
| | 2 | **1.9** | 1.9 | 1.9 | 1.00 | | |
| 189 | 1 | **1.7** | 1.7 | 1.7 | 1.00 | 1.7 | 7.5E-03 |
| | 2 | **1.8** | 1.8 | 1.8 | 1.00 | | |
| 168 | 1 | **1.8** | 1.8 | 1.8 | 0.99 | 1.8 | 8.0E-03 |
| | 2 | **1.7** | 1.7 | 1.7 | 0.99 | | |
| 147 | 1 | **1.5** | 1.4 | 1.5 | 0.99 | 1.5 | 2.2E-03 |
| | 2 | **1.5** | 1.5 | 1.6 | 0.99 | | |
| 126 | 1 | **0.6** | 0.6 | 0.6 | 1.00 | 0.5 | 5.2E-03 |
| | 2 | **0.5** | 0.5 | 0.5 | 0.99 | | |

**Table 2 One way ANOVA for fitted slopes in Table 1 with variance explained with the 126 bp template included and excluded from population. The in group SS (sum of squares) is in the same range as the between group SS for templates 210 to 147 giving a p value >0.05 meaning that the null hypothesis that the slopes of templates 210 to 147 belong to the same population cannot be rejected with above 95 % certainty. In the case where the 126 bp template is included in the ANOVA test the p-value is smaller than 0.05 allowing discrimination of this template from the population under the given test conditions.**

ANOVA, one way

| Source of variance | | SS | df | MS | F | p-value | F-critical |
|---|---|---|---|---|---|---|---|
| Between | 210-126 | 2.24 | 4 | 0.561 | 74.9 | 1.2E-04 | 5.19 |
| groups | 210-147 | 0.11 | 3 | 0.036 | 4.5 | 9.1E-02 | 6.59 |
| In groups | | 0.04 | 5 | 0.007 | | | |
| Total | | 2.28 | 9 | | | | |

The ANOVA sum of squares, *SS,* are the sum of the squared deviations of the estimates from their mean, dividing the in group SS with the degrees freedom of the estimate thus produce the mean variance of 0.007487 $RFU^2/s^2$ for the fitted slopes. Taking the square root of the variance gives the mean standard deviation of 0.086528 RFU/s or about 0.1 RFU/s; we can therefore expect a deviation around 0.1 RFU/s in following experiments. The variance observed comes from the transfer by pipette into the measuring well as well as random instrumental errors as each measurement series represent one transfer. As seen in Table 1 the variance for the circle templates seem to increase with size. This is due to that the template stock concentrations were not normalized prior to loading while all the templates with widely different synthesis efficiencies, Figure 18, in the same elution volume of 50 µ giving much higher concentrations for the larger templates. From this follows that the absolute errors are larger for the same amount template loaded the bigger the size of the template.

## Transcription kinetics of the RNAP T7 on small circular templates

A full range kinetic study was performed on the circular templates of 210 to 147 bp in circumference with the addition of the linear control with the molecular beacon mbAK4P2 as indicator. The study was performed over 11 concentrations, not including the zero, and can be seen with fitted Michaelis-Menten kinetics in Figure 22. The $v_{max}$ value was used rather than the scaled $k_{cat}$ parameter as the enzyme concentration was only expressed in activity units *u* by the supplier. A visual predictive check confirms the validity of the model with the great majority of the measured points falling within the 95% nonsimultaneous observation bounds. The resulting kinetic parameters are presented with 95% confidence interval and $R^2$ goodness of fit measure in Table 3 indicating a generally excellent fit of model to data. Of special notice is the dramatic impact of circularization on the $K_M$ value which increases about one order of from 2.8 nM for the linear control to 19.4 nM for the 210 bp circular template. In general the $v_{max}$ value is much less affected with the maximum speed of transcription dropping from 3.9 to 2.1 mRFU/s upon circularization. A comparative plot of $K_M$ and $v_{max}$ values following circularization and decreasing circular template size is presented in Figure 23. The general trend toward lower $v_{max}$ values for smaller circular templates is broken with the approximate doubling of the $v_{max}$ when going from the 168 to 147 bp template.

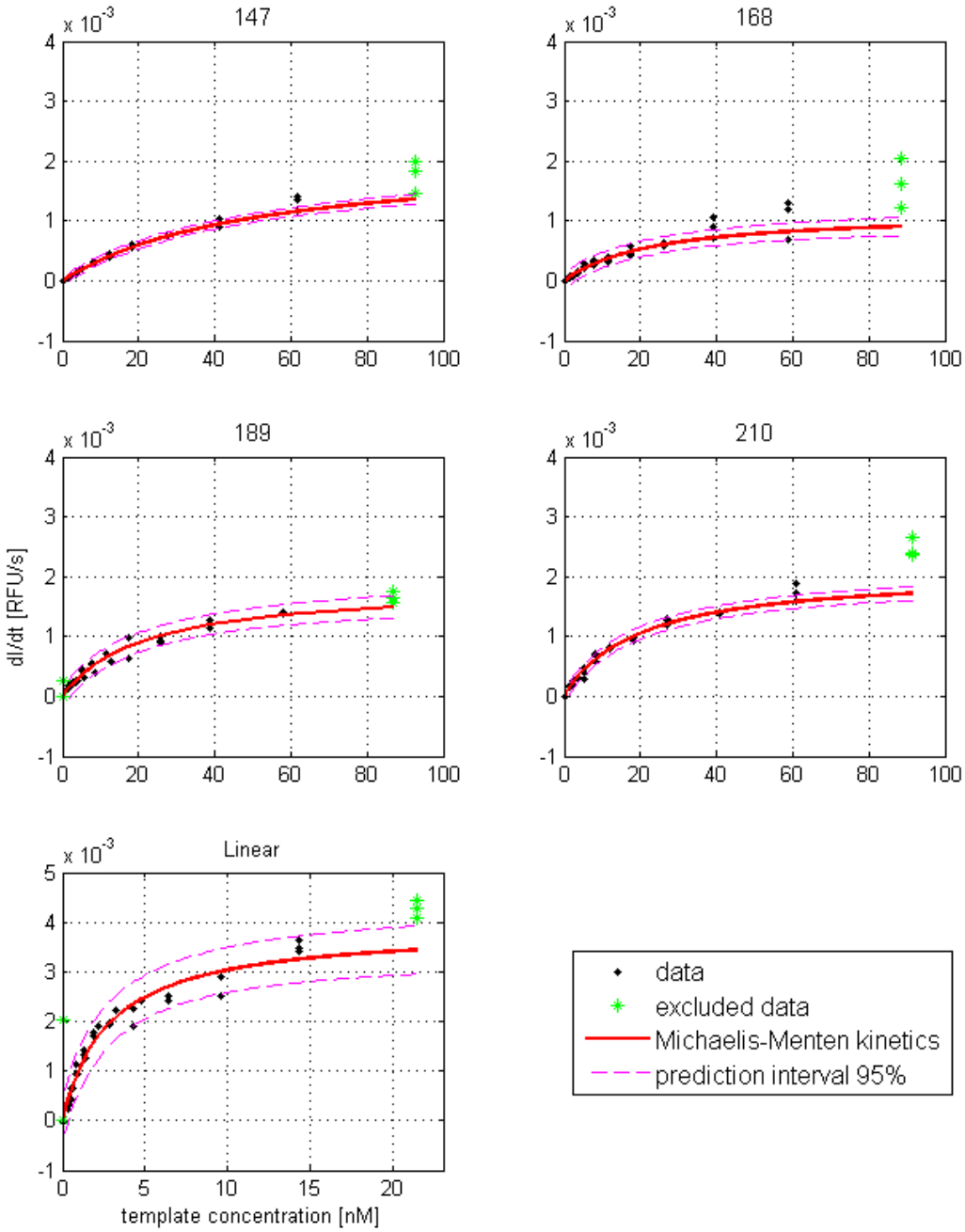



**Figure 22 Plot of transcription kinetics of circular templates from 147 to 210 bp in size and 178 bp linear control. On visual inspection the Michaelis-Menten kinetics seem to fit the recorded data well which is supported by the goodness of fit statistics in Table 3. The data omitted are the loading wells due to a suspected volume error and wells empty due to loading error. Of special notice is the dramatic increase of KM values upon circularisation, possible being indicative of the increased resistance of opening a transcription bubble when acting on a ring closed template.**

**Table 3 Kinetic parameters from fitted Michaelis-Menten kinetics in Figure 22, values of KM and vmax presented with -,+ 95% confidence interval. A sharp drop in KM is observed upon linearization of the template (from linear to 210 bp template) followed by a gentler circle size dependent decrease. The vmax is less severely affected by circularization with the enzyme acting with approximately twice the velocity on the linear template as compared with the circular. The R2 values indicate an overall good fit of model to data but the unexpected increase in vmax from the 168 to 147 bp circle raises questions about experiment precision and/or model validity.**

| size | $K_M$ | | | $v_{max}$ | | | $R^2$ |
|---|---|---|---|---|---|---|---|
| bp | nM | - | + | mRFU/s | - | + | |
| 147 | 50.3 | 42.3 | 58.4 | 2.1 | 1.9 | 2.3 | 0.99 |
| 168 | 23.6 | 16.0 | 31.3 | 1.1 | 1.0 | 1.3 | 0.97 |
| 189 | 21.7 | 16.5 | 26.9 | 1.9 | 1.7 | 2.1 | 0.97 |
| 210 | 19.4 | 17.0 | 21.8 | 2.1 | 2.0 | 2.2 | 0.99 |
| linear | 2.8 | 2.2 | 3.4 | 3.9 | 3.5 | 4.2 | 0.96 |

The very large decrease seen in $K_M$ upon circularization is in good agreement with the idea that the polymerase upon transcription initiation has to overcome the tension put on the template due to opening of an transcription bubble and subsequent over winding of the DNA double helix. While it has been shown that the RNA polymerases can overcome a large torque averse to the transcription this has been under active transcription conditions where the polymerase is part to a large free energy gain from hydroxylation of the NTP: phosphates (Harada, et al. 2001). In contrast the polymerase on transcription initiation has no such source of free energy and has to rely on much smaller energy gains due to configurational changes of the enzyme-template complex (Brieba and Sousa 2001). We can here conclude that while the affinity of the enzyme to its target is decreased the polymerase still maintains a considerable activity on the circular template.

During this study Michaelis-Menten kinetics has been used to characterize and explain the observed system mainly due to the ease of fitting the data to the simple two parameter system. Even though this approach seems to produce a very good correspondence between the observed and predicted values, Figure 22, it serves well to remember that Michaelis-Menten kinetics is a gross simplification of the reality. The binding of the T7 RNAP to its specific promotor, the clearance of the polymerase from the promotor and resulting elongation of the nascent transcript is a complicated multi step process studied in detail elsewhere (Martin, Muller and Coleman 1988) (Rong, He, et al. 1998) (Brieba and Sousa 2001) and is quite out of the scope of this work. However, Michaelis-Menten kinetics *can* under favorable conditions be used to describe complicated, multi step, reactions (Mathews, Holde and Ahren 2000) and *has* been employed to model the T7 RNAP promotor binding and clearance (Martin and Coleman 1987).

The two Michaelis-Menten kinetic s key assumptions, of quasi steady state and constant enzyme concentration (Mathews, Holde and Ahren 2000), are assumed to be met with support of the essentially linear transcription kinetics seen in Figure 21. Also assumed in Michaelis-Menten kinetics

are single substrate reaction conditions which can be met by fixation of all but the studied substrates concentration (Wilson and Walker 2005). In this study this is achieved by providing a gross excess of NTP:s while varying the template concentration. Martin and Coleman have used Michaelis-Menten kinetics to study the binding and release of the RNAP T7 to its promotor (Martin and Coleman 1987) and based on this a model for the transcription of multi target sites from a circular template is proposed:

The standard Michaelis-Menten relation is incorporated as the left hand part of equation (1) with, *E*, *P* and *EP* the enzyme, promotor and enzyme-promotor complex concentrations at steady state. The rate constants $k_1$ and $k_{-1}$ describe the forward and reverse reaction rates for the enzyme-promotor complex formation and incorporates all rate constants up to product formation in the full system. The enzyme-promotor complex *EP* releases the bound promotor *P* and transforms to the elongation complexes *EC* with rate $k_p$, producing the first target $R_1$. The elongation complex *EC* either remains bound to the template or continues one lap around the circle to produce a second target $R_2$ with rate $k_f$ or the enzyme fall off the template and abort transcription with rate $k_{off}$. This is repeated *n* number times until the enzyme releases the template and transcription is aborted.

$$E + P \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} EP \overset{k_p}{\rightarrow} EC + P + R_1 \overset{k_{f1}}{\longrightarrow} \left[ EC + R_i \overset{k_{fi}}{\rightarrow} \right]^n \rightarrow E + R_{n+1} \quad (1)$$
$$\downarrow k_{off1} \quad \downarrow k_{offi}$$
$$E \quad E$$

The relation $k_f/(k_{off} + k_f)$ is thus the fraction of enzymes that remain attached to the template and continues transcription after having produced a target. Under the assumption that the forward and off rates remain constant regardless of the transcript length, the total number of targets, and by association the total length of the transcript, will be given by equation (2)

$$R_{tot} = R_1 + R_2 * \frac{k_f}{k_{off+k_f}} + R_3 * \left(\frac{k_f}{k_{off+k_f}}\right)^2 + \cdots + R_{n+1} * \left(\frac{k_f}{k_{off+k_f}}\right)^n \quad (2)$$

This sum is finite under the condition that $k_f / (k_{off} + k_f) < 1$, that is, there is some falloff during each lap, and the expression on the right of equation (3) is produced. We here assume that each target marker is indistinguishable from the others based on location on the transcript. Under the condition where $k_{off} >> k_f$ the expression collapses back to that for transcription from a linear template, that is a single target *R* is produced.

$$R_{tot} = R * \sum_{n=0}^{\infty} \left(\frac{k_f}{k_{off+k_f}}\right)^n \underset{n \rightarrow \infty}{\overset{k_{off} > 0}{=}} R * \frac{1}{1 - \frac{k_f}{k_{off+k_f}}} \quad (3)$$

The total amount of target produced from a circular template with run around transcription is thus given as a product of the amount produced from template without run-around transcription and a factor dependent on the ratio between falloff and run-around. This factor also affects the monitored production rate of target so that the Michaelis-Menten constant $v_{max}$ is split into $vm_{ax}$ and $(1- k_f /k_{off})^{-1}$ which describe the increase in $v_{max}$ due to run around transcription. The expression (4) then describes the kinetics of transcription from a circular template given that Michaelis-Menten k kinetics is applicable to the circle system.

$$v = \frac{V'_{max}*[P]}{\left(1-\frac{k_f}{k_{off}+k_f}\right)(K_M+[P])} \quad (4)$$

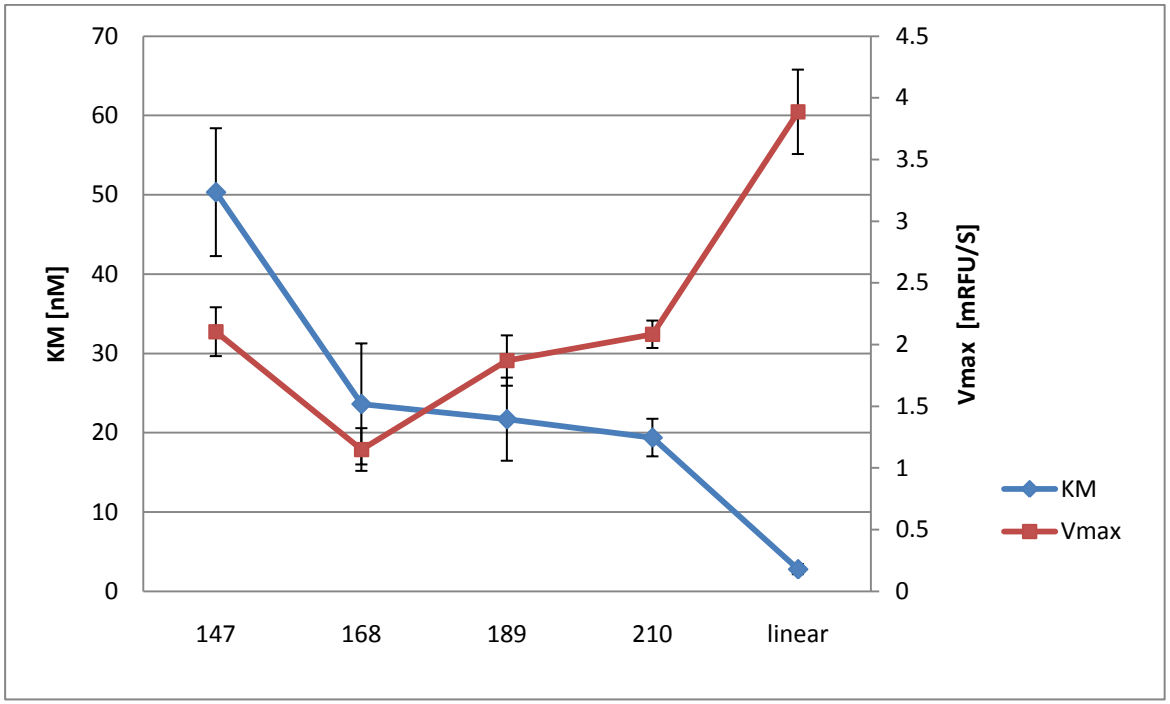



**Figure 23 Plot of fitted Michaelis-Menten kinetic parameters $K_M$ and $v_{max}$ for circular templates of 147 to 210 bp and linear control with error bars for 95% confidence interval shown. Of immediate interest is the general trend of decreasing $v_{max}$ with circle size being broken by the 147 bp template. This could be due to an experimental or model error with an alternative hypothesis being reduced falloff at the double terminator causing increased run around transcription.**

The expression in equation (4) should not be taken as a straight comparison between a linear and a circular template, that is, circularization will *not* automatically increase the transcription rate as the circularization of the substrate may lower both the $K_M$ and $v_{max}$. Indeed as seen in Figure 23 the effects of circularization seem to be a decrease of both these parameters for the range of circles studied in this work, this in spite of the considerable run around transcription observed in Figure 20. Equation (4) does however provide a probable explanation for the increase in observed $v_{max}$ for the 147 bp circular template. The PTH terminator motif incorporated in the template provides the majority of the falloff and the tandem motif is designed to provide in excess of 60% falloff at each occasion in unstrained configuration, see Design: *Circle Family AK4*. The terminator acts in a

configuration dependent manner and deformation of the terminator motif by super coiling has a strong negative impact on the termination efficiency (He, et al. 1998). When the RNA polymerase initiates transcription it unwinds a stretch of the DNA helix approximately one helical turn in length, that is around 11 bp, forming a transcription bubble (Martin and Coleman 1987). The unwinding of a portion of a circular DNA helix causes over winding, or super coiling, of the remaining stretch of DNA, the effect of which becomes greater the smaller the circle. It is thus likely that the increase in $v_{max}$ seen for the 147 bp template is due to lowered termination efficiency as described by equation (4).

Incidentally while the $v_{max}$ sharply increases for the 147 bp template as compared with the preceding templates so does the $K_M$ after having remained relatively unaffected by the decrease in circumference from 210 to 168 bp, Figure 23. In fact the increase in $K_M$ here is dramatic and comparable in absolute, if not in relative, terms with the decrease observed when going from the linear to the circular template. This increase in $K_M$ as compared with the preceding steps is hard to explain purely by an increased resistance to the opening of a transcription bubble as this effect is linear with respect to circle size. A possible answer to this problem is that the increase is seen first after the 168 bp template which is the smallest template estimated as unstrained (Ulanovsky, et al. 1986). The promotor recognition of the T7 is dependent on a number of base pair determinates from which deviations are ill tolerated with respect to initiation efficiency (Rong, He, et al. 1998) . It may be that the strain caused by over bending of circular templates smaller than 168 bp is large enough to disturb the precise configuration of the promotor and thereby radically decrease the enzymes initiation efficiency translating to higher $K_M$. An alternative explanation is that the force required to open and maintain a transcription bubble at the 147 bp circumference is drawing near a threshold of how much force the enzyme is able to exert on the promotor. This type of threshold effect has been observed for RNA polymerases working against an applied force (Yin, et al. 1995).

Top fully confirm or dismiss this hypothesis a study should be performed where super coiling can be ruled out as an contributing factor, e.g. by introducing nicks in the circle permitting rotation of the DNA strand. Furthermore it would be interesting to monitor the transcription kinetics of the 126 bp template which was not included in this study due to lack of template. The single concentration study from the preceding section presented in Table 1 show a much lower slope for the 126 bp template as compared with the 147 bp template at equal concentrations. This result indicates a much lower KM or vmax or both in conjugation.

## Summary and prospects

The results from the transcription studies conclusively show that rolling circle transcription down to 168 bp is functioning with transcripts of length greater than the circle circumference clearly observable on gel, Figure 20. The T7 RNA Polymerase shows transcriptional activity on the 147 and 126 bp templates, Table 1, albeit with markedly decreased rate for the smaller circle. A kinetic study was performed on the 210 to 147 bp circles and a linear control circle fragment, Figure 23, and Michaelis-Menten kinetic parameters fitted, Table 3.The results show a marked increase of $K_M$ by an order of magnitude upon circularization and a simultaneous but lesser decrease of the $v_{max}$. This result explains the difficulty to separate the transcription rates at the single concentration measurement in Table 1. The concentration studied, 100 nM, is much larger than the fitted $K_M$ for the 210 to 147 bp circles, ranging between 20 and 50 nM, and the observed slope is thus close to the

$v_{max}$ for the studied circles which is relatively unaffected by changing circle size,Table 3. Interestingly the transcription rate for the 100 nM 126 bp circle is only about a quarter of that for the 147 to 210 bp circles suggesting a large change in the kinetic parameters.

The finding that the increase in $K_M$ is the main explanation for the decreased transcription rate of the T7 RNAP on circular templates is promising with regard to the integration of the enzyme with DNA nanostructures to create motor complexes. This as the local concentration for the promotor bound in the vicinity of the polymerase would be much greater than the calculated $K_M$. If e. g. the enzyme is bound to a DNA rotaxane of about 150 nm in length with a threaded DNA minicircle containing a T7, promotor the local template concentration for the T7 RNAP will be 167 µM while the $K_M$ for the minicircle is in the nM range. This calculation is based on one promotor restricted to a sphere of a diameter of 75 nm. Thus a polymerase immobilized with a promotor within the scale of a DNA noncomplex can be expected to work with about half the maximum transcription rate on a linear template.

The above assumption should be valid for circles in the 210 to 147 bp size range. For the 126 bp template a major decrease in transcription rate for a 100 nM template concentration as compared with the 147 to 210 bp templates is observed in Table 1. However even if all of the noted decrease would be due to a shift in either of the kinetic parameters the maximum $K_M$ and $v_{max}$ needed to explain the decrease are 300 nM and 0.75 RFU/s respectively. This value for $K_M$ is well below the local concentration achieved in the earlier example while the $v_{max}$ would be about a fifth of that for a linear template, still maintaining considerable transcriptional activity. Even if the $v_{max}$ not only would remain unaffected but continue the increase seen between the 168 and 147 bp templates the resulting $K_M$ would still remain far below the local concentration. We can thus conclude that not only are rolling circle transcription likely to be practical for restricted enzyme template systems of 210 to 147 bp in circumference, it reasonable that it will be practical even with the 126 bp rotaxane system developed by Ackerman *et al* (Ackermann, et al. 2010).

These findings and speculations are all valid under the data observed, it is, however, worth to note that these data are all from single studies of a very limited number of experiments. In particular the kinetic study has flaws which include a trend towards unexpectedly high values for the upper measuring points even after correction for a volume error in the loading well. This ill fits the adapted Michaelis-Menten kinetics and the study should be redone with more care taken in the making of the dilution series and if the effect the model must be reconsidered. Also the position of the measuring points should be reworked with the information at hand to better provide information about the studied systems; ideally this would be done with optimal design software. In addition to above it would be very interesting to perform a kinetic study also for the 126 bp circular template which was left out from the kinetics study in this work due to insufficient template amounts. For all the studied templates a reworked design without the PTH terminator motif would provide more information about the kinetics and stability of the transcription complex, the terminators only necessitated by the demand to be able to monitor the transcription products on a gel. A reworked design utilizing two differently marked molecular beacons against spaced targets on rolling circle transcription product in a multiplex assay would give exact falloff ratios and thus provide information for a more mechanistic model of the rolling circle system.

With the above reservations we can conclude that rolling circle transcription from small circular double stranded DNA templates is possible with moderate influence on the saturation concentrations for the system. This coupled with well proven techniques for coupling of enzymes to DNA (Verma and Eckstein 1998) (Pomerantz, et al. 2005)and the ability to build free floating DNA complexes held together by mechanical bonds developed by the Famulok group of the LIMES institute (Ackermann, et al. 2010) opens the road the construction of functional RNAP based DNA nano motors, Figure 11.

# Materials and methods

## DNA

DNA oligos was ordered from METABION 5' -phosphorylated or unmodified depending on intended usage. The oligos were HPLC purified and verified with mass spectrometry (MS) analysis by the supplier. Sequences and modifications for all oligos used in this study can be found in Appendix 2: AK4 circle family.

## Molecular beacon

2' O-methyl RNA molecular beacon fluorescent probe labelled with 5' fluorophore FAM (6-carboxyfluorescein) and 3' quencher DABCYL (4-(dimethylaminoazo) benzene-4-carboxylic acid) was ordered from Eurogentec. The probe was dual HPLC (RP+RP) purified and quality controlled with MALDI-TOF Mass Spectrometry and Analytical HPLC by the supplier.

## Buffer systems

1 × TAE buffer: 40 mM Tris, 20 mM AcOH, 1mM EDTA.

1 × DNA storage buffer: 10 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$ at pH 7.5.

1 x Transcription buffer: 40 mM Tris-HCL at pH 8, 10 mM NaCl, 10 mM DTT, 6 mM $MgCl_2$, 2 mM Spermidine.

1.25 x Transcription buffer - Mg: 50 mM Tris-HCL at pH 8, 12.5 mM NaCl, 12.5 mM DTT, 2.5 mM Spermidine.

1 x Ligation buffer: 40 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM ATP (pH 7.8 at 25°C).

1 x TBE buffer: 89 mM Tris-borate, 89 mM boric acid, 2mM EDTA.

## HPLC purification

Weak anion exchange HPLC purification: column TSK gel DEAE-NPR4, 6 mm × 35 mm (TOSOH); buffer A: 20 mM Tris-HCl, pH 9.0; buffer B: buffer A plus 1M NaCl; gradient 40% → 65% B in 30min. Following purification samples were concentrated using Ultracel Centrifugal Filter (YM-30, MILLIPORE), washed twice with 1 x DNA storage buffer and eluted in 50 µl 1 x DNA storage buffer.

### Gel electrophoresis

Analytical polyacrylamide (6% PAGE) gels were run in 1 × TAE for 25 min at 200 V, stained with ethidiumbromide and visualized under UV irradiation. Analytical polyacrylamide (9% PAGE) gels were run in 1 × TAE for 40 min at 200 V, stained with ethidiumbromide and visualized under UV irradiation. Denaturing polyacrylamide gels (10% PAGE, 8M Urea) were run in 1 x TBE for 1 h, 170 V, stained with ethidiumbromide and visualized under UV irradiation.

### Synthesis of minicircles and curved fragments

Oligonucleotides (800 pmol) and 40 mM NaCl in 1 x Ligation buffer (200 µl) were annealed from 60 °C to 15 °C over 75 min (dT/dt = 0,6 °C$min^{-1}$). 2 µl (10 u) of T4 DNA ligase (Fermentas) were added and mixture incubated over night at 15 °C. The products were analyzed with analyzed by polyacrylamide gel electrophoresis (6% PAGE in 1 x TAE at 200 V for 25 min). The products were purified by HPLC and concentrated using Ultracel YM-30 Centrifugal Filters.

### Transcription assay

DNA template (0,25 pmol of linear fragment or 1 pmol of circular product) together with 2 mM NTP mix (ATP, UTP, GTP, CTP), 0,6 µl (12 u) T7 RNAP (Fermentas), and 0,2 µl (4 u) RNasin (Promega) In 20 µl 1 x Transcription buffer was incubated at 37 °C for 4 h. The reaction mixture was quenched by the addition of 1 µl EDTA and the transcription products analyzed by denaturing gel electrophoresis (10% PAGE with 8M Urea in 1 x TBE at 200 v for 45 min).

### Real time point assay

DNA template (125 nM) together with 2 mM NTP mix (ATP, UTP, GTP, CTP), 0.75 u/µl T7 RNAP (Fermentas), 0.5 u/µl RNasin (Promega ), and 250 nM molecular beacon in 40 µl 1.25 x Transcription buffer -Mg was put in 96-well half area flat bottom fluorescence reader plates. The reaction was started by the injection of 10 µl 30 mM $MgCl_2$ (6 mM final concentration) for a final volume of 50 µl per well and monitored on a Thermo Electron Varioskan 3.01.15 fluorescence plate reader, at temperature 37 °C, with 30 s intervals at 120 measurement points, detection wavelength 517 nm, excitation wavelength 491 nm, and excitation bandwidth 12 nm.

The data was analyzed using an iterative linear regression method for finding the linear range of the florescence readout from each well. The equation for the straight line was fit to the whole data set and a 95% observation simultaneous symmetric prediction interval was calculated. Observed values outside this interval were excluded from the data set and the linear regression for the new subset was calculated. This process was iteratively repeated until the change in slope between two steps was less than 1% and the resulting slope value was assumed to be the steady state transcription rate for the studied concentration.

### Real-Time transcription kinetic assay

DNA template together with 2 mM NTP mix (ATP, UTP, GTP, CTP), 0.75 u/µl T7 RNAP (Fermentas), 0.5 u/µl RNasin (Promega ), and 250 nM molecular beacon in 120 µl 1.25 x Transcription buffer -Mg was put in the first row of 96-well half area flat bottom fluorescence reader plates. A dilution series over 11 steps in 2/3 intervals from a final concentration of 20 nM for the linear fragment and 100 nM for

circular products was set up with the twelfth well at zero template presence. The concentration was set for the first well by transfer of the calculated amount template to a total of 120 µl loading volume. The dilution series was then set up by transfer of 80 µl of this loading solution to 40 µl of clean transcription buffer in the following well and repeating till the end of the series. The reaction was started by the injection of 10 µl 30 mM $MgCl_2$ (6 mM final concentration) for a final volume of 50 µl per well and monitored on a Thermo Electron Varioskan 3.01.15 fluorescence plate reader, at temperature 37 °C, with 30 s intervals at 120 measurement points, detection wavelength 517 nm, excitation wavelength 491 nm, and excitation bandwidth 12 nm.

The data was analyzed using an iterative linear regression method for finding the linear range of the florescence readout from each well. The equation for the straight line was fit to the whole data set and a 95% observation simultaneous symmetric prediction interval was calculated. Observed values outside this interval were excluded from the data set and the linear regression for the new subset was calculated. This process was iteratively repeated until the change in slope between two steps was less than 1% and the resulting slope value was assumed to be the steady state transcription rate for the studied concentration.

The Michaelis-Menten equation was fitted to the slope values of the linear ranges for the dilution series giving the kinetic parameters $K_M$ and $v_{max}$. The slope from the loading well was excluded from the regression analysis with the following rationale:

The loading well includes a systematic volume error relative to the loading errors for the following wells of the dilution series that can be assumed to be random with a mean of zero. The concentration in the loading well, $c_0$ is:

$$c_0 = \frac{n}{3V + \varepsilon}$$

Where n is the amount template loaded, V the final volume in the well and $\varepsilon$ the volume error in the first well. Two thirds of the volume in the loading well is removed and added to the one volume unit in the next well which is then repeated for the full dilution series. This gives that well *i* of the i:th position in the dilution series has the final concentration:

$$c_i = \frac{{}^{2}/_{3}{}^{i} * n}{3V + \varepsilon}$$

Which for the loading well at position *0* in the dilution series collapses back to equation: , the concentrations are thus proportional for all position in the dilution series regardless of the initial volume error in the loading well. When considering the actual amounts of template, $n_i$, in the wells the loading well differs by the error factor $\varepsilon$:

$$n_0 = \frac{(V + \varepsilon) * n}{3V + \varepsilon}$$

While the remaining wells i = 1 .. 12 follow:

$$n_i = \frac{V * {}^{2}/_{3}{}^{i} * n}{3V + \varepsilon}$$

The fluorescent readout from the loading well will thus be off from the expected proportional readout from the dilution series by the unknown error factor$\varepsilon$. To avoid influence of the loading error the first data point in all series is excluded.

## Acknowledgments

I wish to express my gratitude to:

Professor Michael Famulok for granting me access to the resources of his lab making this thesis possible. Doktor Damian Ackerman who graciously took the time to share his expertise on the subject and to give me advice on my work. All the nice people at the Laboratory of Chemical Biology at Universität Bonn whom gave valuable input and provided invaluable chit-chat –Jeff, thanks for all the quips!

Professor N. C. Seeman, Professor Ponzy Lu, Professor Andrew J Turberfield, Dr. Damian Ackerman, Professor William McAllister and Professor Mats Nilsson for graciously granting me figures for the illustration of this work.

The Erasmus lifelong learning program which covered the costs of my stay in Germany. Of the Uppsala University I deeply thank my programme coordinators Margareta Krabbe, Torgny Fornstedt and Lars-Göran Josefsson for being at hand and getting all the practicalities sorted on short notice

My opponents Olof Jönsson and Gunnar Dahlberg whom provided valuable input.

My parents Jan and Marie who have (most of the time) encouraged my curiosity and my grandmother Alice – Thank you for all the hours at the coffee table listening to me going on about whatever subject!

Sandra my love, you have had great patience with me.

Lastly I much appreciate the work of my scientific reviewer Professor Gerhart Wagner – thank you for accepting to review this work and for given advice.

# References


Ackermann, Damian, Thorsten L Schmidt, Jeffrey S Hannam, Chandra S Purohit, Alexander Heckel, and Michael Famulok. “A double-stranded DNA rotaxane.” *Nat Nano* 5 (06 2010): 436--442.

Andronescu, Mirela, Anne Condon, Holger H Hoos, David H Mathews, and Kevin P Murphy. “Efficient parameter estimation for RNA secondary structure prediction.” *Bioinformatics* 23 (7 2007): 19--28.

Bernhart, Stephan, Hakim Tafer, Ulrike Muckstein, Christoph Flamm, Peter Stadler, and Ivo Hofacker. “Partition function and base pairing probabilities of RNA heterodimers.” *Algorithms for Molecular Biology* 1 (2006): 3.

Böhm, Konrad J, Roland Stracke, Peter Mühlig, and Eberhard Unger. “Motor protein-driven unidirectional transport of micrometer-sized cargoes across isopolar microtubule arrays.” *Nanotechnology* 12 (2001): 238--244.

Bravo, José A., Françisco M. Raymo, Stoddart J. Fraser, Andrew J. P. White, and David J. Williams. “High Yielding Template-Directed Syntheses of [2] Rotaxanes.” *European Journal of Organic Chemistry* 11 (1998): 2565--2571.

Breaker, Ronald R, Amrita Banerj, and Gerald F. Joyce. “Continuous in Vitro Evolution of Bacteriophage RNA Polymerase Promoters.” *Biochemistry* 33 (1994): 11980-11986.

Breaker, Ronald R, and Gerald F Joyce. “A DNA enzyme that cleaves RNA.” *Chemistry & Biology* 1 (12 1994): 223--229.

Breaker, Ronald R., Amrita Banerji, and Gerald F. Joyce. “Continuous in Vitro Evolution of Bacteriophage RNA Polymerase Promoters.” *Biochemistry* 33 (1994): 11980--11986.

Brieba, Luis G, and Rui Sousa. “T7 promoter release mediated by DNA scrunching.” *EMBO J* (European Molecular Biology Organization) 20 (12 2001): 6826--6835.

Cazenave, C, and O C Uhlenbeck. “RNA template-directed RNA synthesis by T7 RNA polymerase.” *Proceedings of the National Academy of Sciences of the United States of America* 91 (07 1994): 6972--6976.

Cohen, Stanley N, Annie C Chang, Herbert W Boyer, and Robert B Helling. “Construction of Biologically Functional Bacterial Plasmids In Vitro.” *Proceedings of the National Academy of Sciences of the United States of America* 70 (1973): 3240--3244.

Daube, SS, and PH von Hippel. “Functional transcription elongation complexes from synthetic RNA-DNA bubble duplexes.” *Science* 258, no. 5086 (11 1992): 1320--1324.

Diaz, George A., Curtis A. Raskin, and William T. McAllister. “Hierarchy of Base-Preference in the Binding Domain of the Bacteriophage T7 Promoter.” *Journal of Molecular Biology* 229, no. 4 (1993): 805--811.

Diekmann, Stephan, and James C Wang. “On the sequence determinants and flexibility of the kinetoplast DNA fragment with abnormal gel electrophoretic mobilities.” *Journal of Molecular Biology* 186 (1985): 1--11.

Dunn, John J., F. William Studier, and M. Gottesman. “Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements.” *Journal of Molecular Biology* 166, no. 4 (1983): 457--689.

Egholm, Michael, et al. “PNA hybridizes to complementary oligonucleotides obeying the Watson&Crick hydrogen-bonding rules.” *Nature* 365 (10 1993): 566--568.

Gruber, Andreas R, Ronny Lorenz, Stephan H Bernhart, Richard Neubock, and Ivo L Hofacker. “The Vienna RNA Websuite.” *Nucl. Acids Res.*, 4 2008: W70-W74.

Gu, Hongzhou, Jie Chao, Shou-Jun Xiao, and Nadrian C. Seeman. “A proximity-based programmable DNA nano scale assembly line.” 465 (May 2010): 202--207.

Hagerman, Paul J. ”Flexibility of DNA.” *Annual Review of Biophysics and Biophysical Chemistry* (Annual Reviews) 17 (06 1988): 265--286.

Han, Wenhai, S M Lindsay, Mensur Dlakic, and Rodney E Harrington. “Kinked DNA.” *Nature* 386 (04 1997): 563--563.

Harada, Yoshie, Osamu Ohara, Akira Takatsuki, Hiroyasu Itoh, Nobuo Shimamoto, and Kazuhiko, Kinosita. “Direct observation of DNA rotation during transcription by Escherichia.” *Nature* 409 (01 2001): 113--115.

He, Biao, et al. “Characterization of an Unusual, Sequence-specific Termination Signal for T7 RNA Polymerase.” *THE JOURNAL OF BIOLOGICAL CHEMISTRY* 273, no. 30 (July 1998): 18802--18811.

Hofacker, I L, W Fontana, P F Stadler, L S Bonhoeffer, M Tacker, and P Schuster. “Fast folding and comparison of RNA secondary structures.” *Monatshefte für Chemie / Chemical Monthly* 125 (02 1994): 167--188.

Iwata, Masaaki, Masaki Izawa, Nobuya Sasaki, Yoko Nagumo, Hiroyuki Sasabe, and Yoshihide Hayashizaki. “T7 RNA polymerase activation and improvement of the transcriptional sequencing by polyamines.” *Bioorganic & Medicinal Chemistry* 8 (8 2000): 2185--2194.

Jia, Lili, Samira G Moorjani, Thomas N Jackson, and William O Hancock. “Microscale Transport and Sorting by Kinesin Molecular Motors.” *Biomedical Microdevices* 6 (03 2004): 67--74.

Kaur, Harleen, Amit Arora, Jesper Wengel, and Souvik Maiti. “Thermodynamic, Counterion, and Hydration Effects for the Incorporation of Locked Nucleic Acid Nucleotides into DNA Duplexes.” *Biochemistry* (American Chemical Society) 45 (06 2006): 7347--7355.

Ke, Yonggang, Jaswinder Sharma, Minghui Liu, Kasper Jahn, Yan Liu, and Hao Yan. “Scaffolded DNA Origami of a DNA Tetrahedron Molecular Container.” *Nano Letters* (American Chemical Society) 9 (06 2009): 2445--2447.

Liang, Xingguo, Hidenori Nishioka, Nobutaka Takenaka, and Hiroyuki Asanuma. "A DNA Nanomachine Powered by Light Irradiation." 9, no. 5 (02 2008): 702--705.

Limberis, Loren, and Russell J Stewart. “Toward kinesin-powered microdevices.” *Nanotechnology* 11 (2000): 47--51.

Lindahl, Tomas. ”Instability and decay of the primary structure of DNA.” *Nature* 362 (04 1993): 709--715.

Liu, Dage, Mingsheng Wang, Zhaoxiang Deng, Richard Walulu, and Chengde Mao. “Tensegrity: a Construction of Rigid DNA Triangles with Flexible Four-Arm DNA Junctions.” *Journal of the American Chemical Society* (American Chemical Society) 126 (03 2004): 2324--2325.

Liu, Haiqing, et al. “Control of a biomolecular motor-powered nanodevice with an engineered chemical switch.” *Nat Mater* 1 (11 2002): 173--177.

Liu, Jianwei, and Patricia Feldman. “Real-Time Monitoring in Vitro Transcription Using Molecular Beacons.” *Analytical Biochemistry* 300 (2002): 40--45.

Liu, Mingzhe, Hiroyuki Asanuma, and Makoto Komiyama. “Azobenzene-Tethered T7 Promoter for Efficient Photoregulation of Transcription.” *Journal of the American Chemical Society* (American Chemical Society) 128 (01 2006): 1009--1015.

Lund, Kyle, et al. “Molecular robots guided by prescriptive landscapes.” *Nature* 465 (May 2010): 206–210.

MacDonald, Douglas, Kristina Herbert, Xiaolin Zhang, Thomas Polgruto, and Ponzy Lu. “Solution Structure of an A-tract DNA Bend.” *Journal of Molecular Biology* 306 (2001): 1081--1098.

Mackay, John, and Olfert Landt. “Real-Time PCR Fluorescent Chemistries.” 2007. 237--261.

Marini, Joan C, Stephen D Levene, Donald M Crothers, and Paul T Englund. “Bent helical structure in kinetoplast DNA.” *Proceedings of the National Academy of Sciences of the United States of America* 79 (1982): 7664-7668.

Marras, Salvatore A, Benjamin Gold, Fred R Kramer, Issar Smith, and Sanjay Tyagi. “Real-time measurement of in vitro transcription.” *Nucl. Acids Res.* 32 (2004): e72.

Marras, Salvatore A.E., Sanjay Tyagi, and Fred Russell Kramer. “Real-time assays with molecular beacons and other fluorescent nucleic acid hybridization probes.” *Clin Chim Acta* 363 (09 2006): 48--60.

Martin, Craig T, Daniel K Muller, and Joseph E Coleman. “Processivity in early stages of transcription by T7 RNA polymerase.” *Biochemistry* 27 (1988): 3966--3974.

Martin, Craig T., and Joseph E. Coleman. “Kinetic Analysis of T7 RNA Polymerase-Promoter Interactions with Small Synthetic Promotors.” *Biochemistry* 26 (1987): 2690--2696.

Mathews, Christopher K., K. E. van Holde, and Kevin G. Ahren. “The Kinetics of Enzyme Catalysis.” In *Biochemistry*, by Christopher K. Mathews, K. E. van Holde and Kevin G. Ahren, 375--383. Benjamin/Cummings, 2000.

Mayer, Günter, Damian Ackermann, Nicole Kuhn, and Michael Famulok. “Construction of DNA Architectures with RNA Hairpins.” *Angewandte Chemie International Edition* 47, no. 5 (January 2008): 971--973.

Nacheva, Genoveva A., and Alfredo Berzal-Herranz. “Preventing non desired RNA-primed RNA extension catalyzed by T7 RNA polymerase.” *European Journal of Biochemistry* 270 (2003): 1458--1465.

Nilsson, Mats, Mats Gullberg, Fredrik Dahl, Karoly Szuhai, and Anton K Raap. “Real-time monitoring of rolling-circle amplification using a modified molecular beacon design.” *Nucl. Acids Res.* 30 (2002): e66.

Peffer, N J, et al. “Strand-invasion of duplex DNA by peptide nucleic acid oligomers.” *Proceedings of the National Academy of Sciences of the United States of America* 90 (1993): 10648--10652.

Pomerantz, Richard T, et al. “A Tightly Regulated Molecular Motor Based upon T7 RNA Polymerase.” *Nano Letters* (American Chemical Society) 5 (09 2005): 1698--1703.

Qiu, Hangxia, John C Dewan, and Nadrian C Seeman. “A DNA decamer with a sticky end: the crystal structure of d-CGACGATCGT.” *Journal of Molecular Biology* 267 (4 1997): 881--898.

Rasched, Goran, Damian Ackermann, Thorsten L. Schmidt, Peter Broekmann, Alexander Heckel, and Michael Famulok. “DNA Minicircles with Gaps for Versatile Functionalization.” *Angewandte Chemie International Edition* 47, no. 5 (December 2007): 967--970.

Robertson, Keith D, and Alan P Wolffe. “DNA methylation in health and disease.” *Nat Rev Genet* 1 (10 2000): 11--19.

Rong, Minqing, Biao He, William T McAllister, and Russell K Durbin. “Promoter specificity determinants of T7 RNA polymerase.” *Proceedings of the National Academy of Sciences of the United States of America* 95 (1998): 515--519.

Rong, Minqing, Russell K Durbin, and William T McAllister. “Template Strand Switching by T7 RNA Polymerase.” *Journal of Biological Chemistry* 273 (04 1998): 10253--10260.

Rothemund, Paul W. “Folding DNA to create nanoscale shapes and patterns.” *Nature* 440 (03 2006): 297--302.

SantaLucia, John. ”A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbour thermodynamics.” *Proceedings of the National Academy of Sciences of the United States of America* 95 (02 1998): 1460--1465.

Seeman, Nadrian C. ”DNA in a material world.” *Nature* 421 (01 2003): 427--431.

Seeman, Nadrian C, and Philip S Lukeman. “Nucleic acid nanostructures: bottom-up control of geometry on the nanoscale.” *Reports on Progress in Physics* 68 (2005): 237--270.

Seeman, Nadrian C. ”Nucleic Acid Junctions and Lattices.” *Journal of Theoretical Biology* 99 (1982): 237--247.

Shen, Zhiyong, Hao Yan, Tong Wang, and Nadrian C Seeman. “Paranemic Crossover DNA: A Generalized Holliday Structure with Applications in Nanotechnology.” *Journal of the American Chemical Society* (American Chemical Society) 126 (02 2004): 1666--1674.

Slotkin, R K, and Robert Martienssen. “Transposable elements and the epigenetic regulation of the genome.” *Nat Rev Genet* 8 (04 2007): 272--285.

Soong, Ricky K, George D Bachand, Hercules P Neves, Anatoli G Olkhovets, Harold G Craighead, and Carlo D Montemagno. “Powering an Inorganic Nanodevice with a Biomolecular Motor.” *Science* 290 (11 2000): 1555--1558.

Sproat, B, S, A, I Lamond, B Beijer, P Neuner, and U Ryder. “Highly efficient chemical synthesis of 2'-O-methyloligoribonucleotides and tetrabiotinylated derivatives; novel probes that are resistant to degradation by RNA or DNA specific nucleases.” *Nucleic Acids Research.* 17, no. 9 (May 1989): 3373--3386.

Stahl, F W. ”The Holliday Junction on Its Thirtieth Anniversary.” *Genetics* 138 (10 1994): 241--246.

Tsourkas, Andrew, Mark A Behlke, and Gang Bao. “Hybridization of 2'-O-methyl and 2'-deoxy molecular beacons to RNA and DNA targets.” *Nucl. Acids Res.* 30 (12 2002): 5168--5174.

Tyagi, Sanjay, and Fred R Kramer. “Molecular Beacons: Probes that Fluoresce upon Hybridization.” *Nat Biotech* 14 (03 1996): 303--308.

Tyagi, Sanjay, Diana P Bratu, and Fred R Kramer. “Multicolor molecular beacons for allele discrimination.” *Nat Biotech* 16 (01 1998): 49--53.

Ulanovsky, Levy, Mordechai Bodner, Edward N. Trifonov, and Mordechai Choder. “Curved DNA: Design, synthesis, and circularization.” *Proceedings of the National Academy of Sciences* 83 (1986): 862--866.

Unrau, Hani S., and Peter J. Zaher. “T7 RNA Polymerase Mediates Fast Promoter-Independent Extension of Unstable Nucleic Acid Complexes.” *Biochemistry* 43 (2004): 7873--7880.

Verma, Sandeep, and Fritz Eckstein. “MODIFIED OLIGONUCLEOTIDES: Synthesis and Strategy for Users.” *Annual Review of Biochemistry* (Annual Reviews) 67 (07 1998): 99--134.

Vet, Jacqueline A, and Salvatore A Marras. “Design and Optimization of Molecular Beacon Real-Time Polymerase Chain Reaction Assays.” Chap. 17 in *Methods in Molecular Biology: Oligonucleotide Synthesis*, 273--290. 2005.

Wang, Yinli, John E. Mueller, Borries Kemper, and Nadrian C. Seeman. “Assembly and Characterization of Five-Arm and Six-Arm DNA Branched Junctions.” *Biochemistry* 30 (1991): 5667--5674.

Watson, J D, and F H Crick. “Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid.” *Nature* 171 (04 1953): 737--738.

Wheeler, Richard (Zephyris). “Holliday junction coloured.” *Wikimedia Commons.* 03 02 2007. http://commons.wikimedia.org/wiki/File:Holliday_junction_coloured.png (accessed 10 10, 2010).

Wilson, Kieth, and John Walker. “Enzyme steady state kinetics.” In *Principles and Techniques of Biochemistry and Molecular Biology*, by Kieth Wilson and John Walker, 679--702. Cambridge University Press, 2005.

Yagil, Gad. ”Paranemic Structures of DNA and their Role in DNA Unwinding.” *Critical Reviews in Biochemistry and Molecular Biology* (Informa Healthcare) 26 (01 1991): 475--559.

Yakovchuk, Peter, Ekaterina Protozanova, and Maxim D Frank-Kamenetskii. “Base-stacking and base-pairing contributions into thermal stability of the DNA double helix.” *Nucl. Acids Res.* 34 (1 2006): 564--574.

Yin, Hong, Michelle D Wang, Karel Svoboda, Robert Landick, Steven M Block, and Jeff Gelles. “Transcription Against an Applied Force.” *Science* 270 (1995): 1653--1657.

Yin, Peng, Harry M Choi, Colby R Calvert, and Niles A Pierce. “Programming biomolecular self-assembly pathways.” *Nature* (Nature Publishing Group) 451 (01 2008): 318--322.

Yu, Hua, Kyubong Jo, Kristy L Kounovsky, Juan J Pablo, and David C Schwartz. “Molecular Propulsion: Chemical Sensing and Chemotaxis of DNA Driven by RNA Polymerase.” *Journal of the American Chemical Society* (American Chemical Society) 131 (04 2009): 5722--5723.

Yurke, Bernard, Andrew J Turberfield, Allen P Mills, Friedrich C Simmel, and Jennifer L Neumann. “A DNA-fuelled molecular machine made of DNA.” *Nature* (Macmillian Magazines Ltd.) 406 (08 2000): 605--608.

Zheng, Jianping, et al. “From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal.” *Nature* 461 (09 2009): 74--77.

Zheng, Jiwen, Pamela E Constantinou, Christine Micheel, A P Alivisatos, Richard A Kiehl, and Nadrian C Seeman. “Two-Dimensional Nanoparticle Arrays Show the Organizational Power of Robust DNA Motifs.” *Nano Letters* 6 (07 2006): 1502--1504.

## Appendix 1: Distance elements.

Distance element selected from all available SNNS sequences on the basis of high melting temperature calculated using SantaLucia[3] nearest-neighbour thermodynamics padded with 5xA and checked against dimer formation by sequence alignment in set utilizing an NUC44 scoring matrix. Starting set of 256 sequences, 64 left after selection. Sequences given in 5' to 3' order.

| | | | |
|---|---|---|---|
| 'CACCC' | 'GACCC' | 'CGACC' | 'GGACC' |
| 'CACGG' | 'GACGG' | 'CGAGG' | 'GGAGG' |
| 'CAGCC' | 'GAGCC' | 'CGCAC' | 'GGCAC' |
| 'CAGGG' | 'GAGGG' | 'CGCCG' | 'GGCCG' |
| 'CCACC' | 'GCACC' | 'CGCGC' | 'GGCGC' |
| 'CCAGG' | 'GCAGG' | 'CGCTG' | GGCTG' |
| 'CCCAC' | 'GCCAC' | 'CGGAC' | 'GGGAC' |
| 'CCCCG' | 'GCCCG' | 'CGGCG' | 'GGGCG' |
| 'CCCGC' | 'GCCGC' | 'CGGGC' | 'GGGGC' |
| 'CCCTG' | 'GCCTG' | 'CGGTG' | 'GGGTG' |
| 'CCGAC' | 'GCGAC' | 'CGTCC' | 'GGTCC' |
| 'CCGCG' | 'GCGCG' | 'CGTGG' | 'GGTGG' |
| 'CCGGC' | 'GCGGC' | 'CTCCC' | 'GTCCC' |
| 'CCGTG' | 'GCGTG' | 'CTCGG' | 'GTCGG' |
| 'CCTCC' | 'GCTCC' | 'CTGCC' | 'GTGCC' |
| 'CCTGG' | 'GCTGG' | 'CTGGG' | 'GTGGG' |

[3] SantaLucia, John. "A unified view of polymer, dumbbell, and oligonucleotide DNA nearest -

## Appendix 2: AK4 circle family

Modular circle family consisting of five minicircles, 210 – 105 bp, and two blunt end circle fragments made from 16 oligos. All circles have identical (primary) transcripts.

Faulty promotor sequence corrected, uses substitute segement AK4P2 for segement AK4P.

New promotor element for AK4 family:

```
-20   -1 +6
 AAT TAATACGACTCACTATA GGGAGA
```

I will use the -22 to -18 element GAAAT, the -17 to -1 consensus promotor sequence T AAT ACG ACT CAC TAT A and the consensus initiator sequence GGGAGA from position +1 to + 6. Based on the work of Breaker et al (Breaker, Banerj and Joyce, Continuous in Vitro Evolution of Bacteriophage RNA Polymerase Promoters 1994) I will truncate the sequence element previous to the consensus sequence to AAT also in correspondence with the work by Martin and Coleman (Martin and Coleman 1987)

Promotor consensus:

```
   +6 +1   -20
CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCT
GACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGA
********************iiiiiippppppppppppppppprrr
```

## AK410

All fragments full length, 210 bp

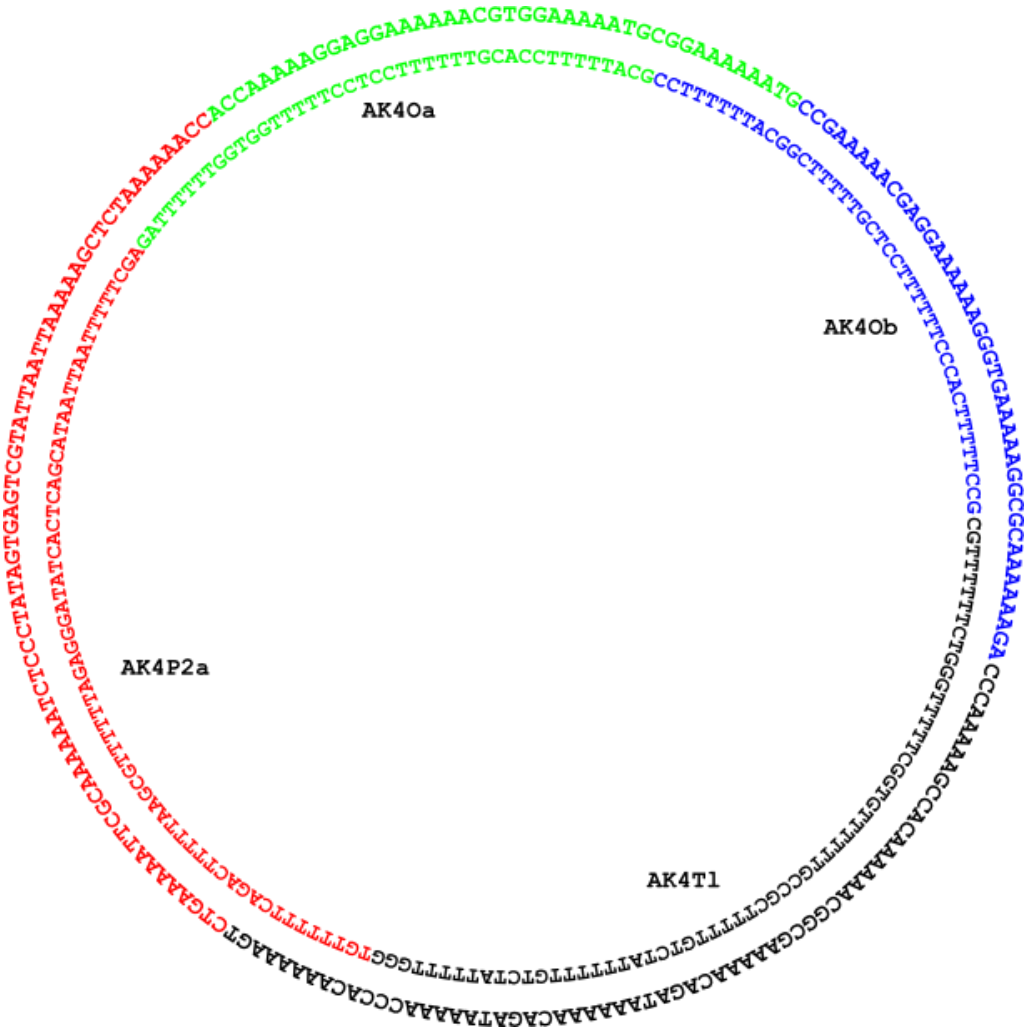
```
AK4T1
  CCCAAAAAGCCACAAAAAACGGCGAAAAACAGATAAAAAACAGATAAAAACCCACAAAAAAGT
CGTTTTTTCTGGGTTTTTCGGTGTTTTTTGCCGCTTTTTGTCTATTTTTTGTCTATTTTTGGG
     --------t --------t********

AK4P2a

  CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTCTAAAAAACC
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGA
******************************.....<_--------------prrr

AK4Oa
  ACCAAAAAGGAGGAAAAAACGTGGAAAAATGCGGAAAAAATG
GATTTTTTGGTGGTTTTTCCTCCTTTTTTGCACCTTTTTACG

AK4Ob
  CCGAAAAACGAGGAAAAAAGGGTGAAAAAGGCGCAAAAAAGA
CCTTTTTTACGGCTTTTTGCTCCTTTTTTCCCACTTTTTCCG
```



## AK49

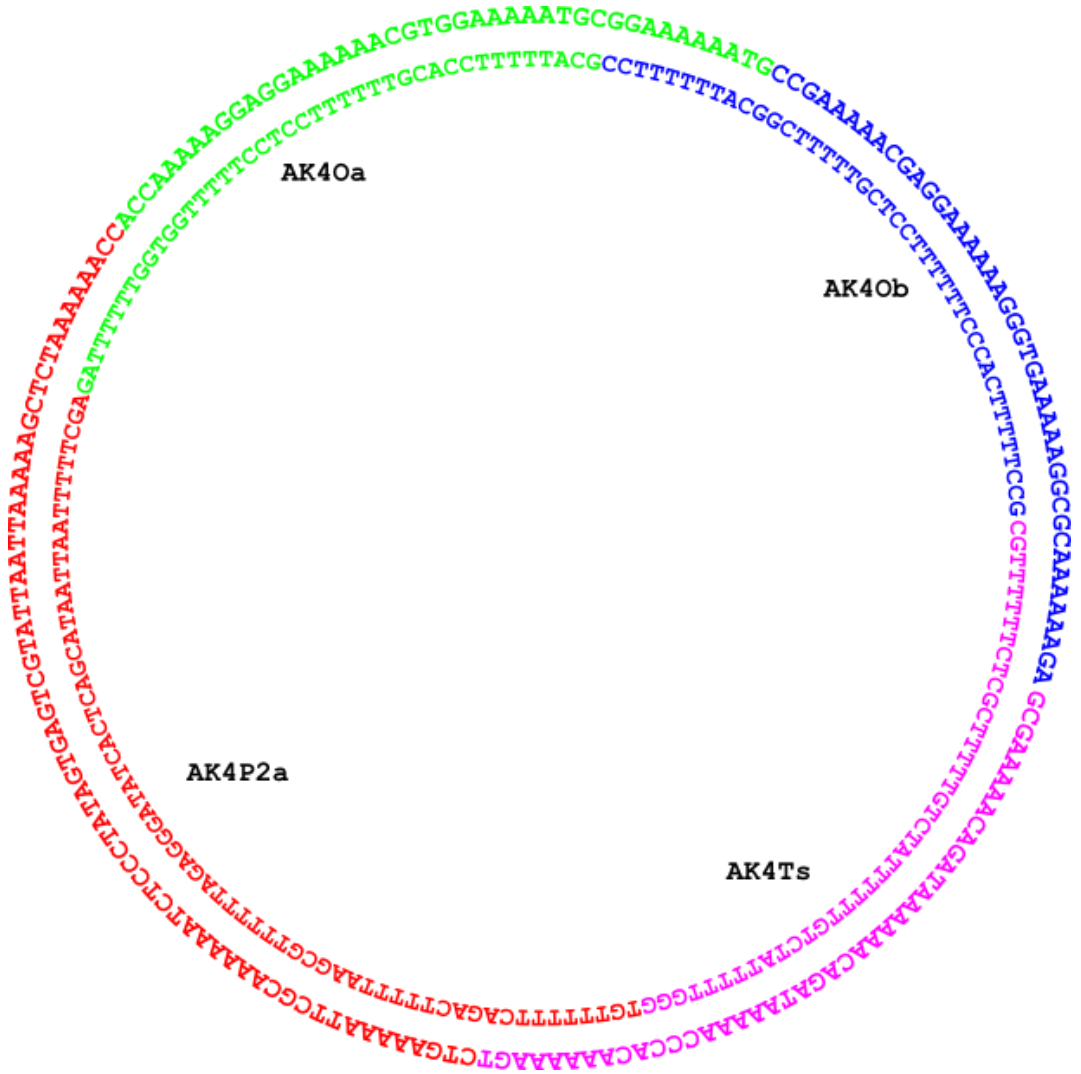
Short termination segment AK4Ts, 189 bp

```
AK4Ts
  GCGAAAAACAGATAAAAAACAGATAAAAACCCACAAAAAAGT
CGTTTTTTCTCGCTTTTTGTCTATTTTTTGTCTATTTTTGGG
  --------t --------t********
AK4P2a

  CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTCTAAAAAACC
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGA
*******************************.....<_--------------prrr

AK4Oa
  ACCAAAAAGGAGGAAAAAACGTGGAAAAATGCGGAAAAAATG
GATTTTTTGGTGGTTTTTCCTCCTTTTTTGCACCTTTTTACG

AK4Ob
  CCGAAAAACGAGGAAAAAAGGGTGAAAAAGGCGCAAAAAAGA
CCTTTTTTACGGCTTTTTGCTCCTTTTTTCCCACTTTTTCCG
```

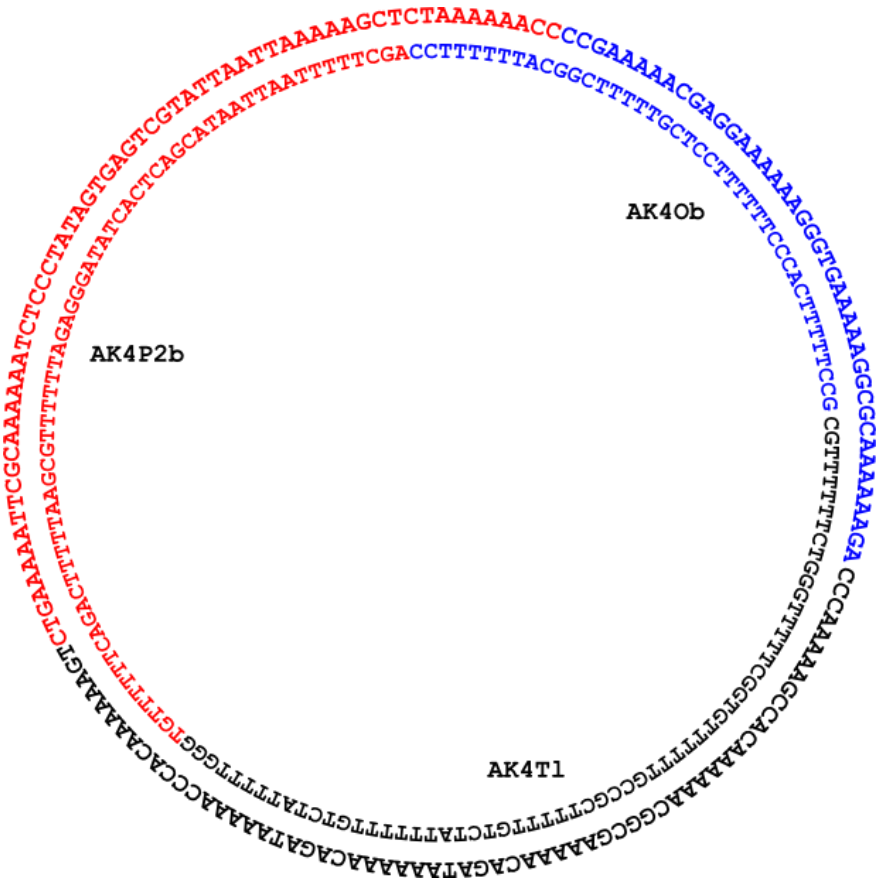


## AK48

Promotor segment AK4Pb, removed fragment AK4Oa, 168 bp

```
AK4T1
  CCCAAAAAGCCACAAAAAACGGCGAAAAACAGATAAAAAACAGATAAAAACCCACAAAAAAGT
CGTTTTTTCTGGGTTTTTCGGTGTTTTTTGCCGCTTTTTGTCTATTTTTTGTCTATTTTTGGG
     --------t --------t********
AK4P2b
  CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTGGAAAAAATG
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGA
******************************.....<--------------prrr

AK4Ob
  CCGAAAAACGAGGAAAAAAGGGTGAAAAAGGCGCAAAAAAGA
CCTTTTTTACGGCTTTTTGCTCCTTTTTTCCCACTTTTTCCG
```

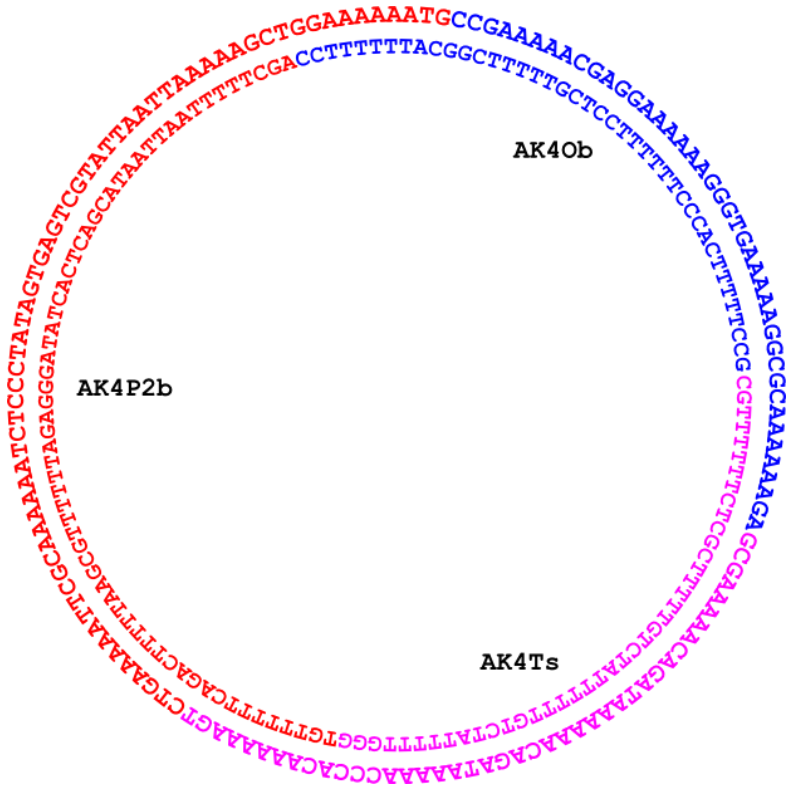


## AK47

Short termination segment AK4Ts, promotor segment AK4Pb, removed fragment AK4Oa, 147 bp

```
AK4Ts
  GCGAAAAACAGATAAAAAACAGATAAAAACCCACAAAAAAGT
CGTTTTTTCTCGCTTTTTGTCTATTTTTTGTCTATTTTTGGG
  --------t --------t********
AK4P2b
  CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTGGAAAAAATG
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGA
******************************.....<--------------prrr

AK4Ob
  CCGAAAAACGAGGAAAAAAGGGTGAAAAAGGCGCAAAAAAGA
CCTTTTTTACGGCTTTTTGCTCCTTTTTTCCCACTTTTTCCG
```



## AK46

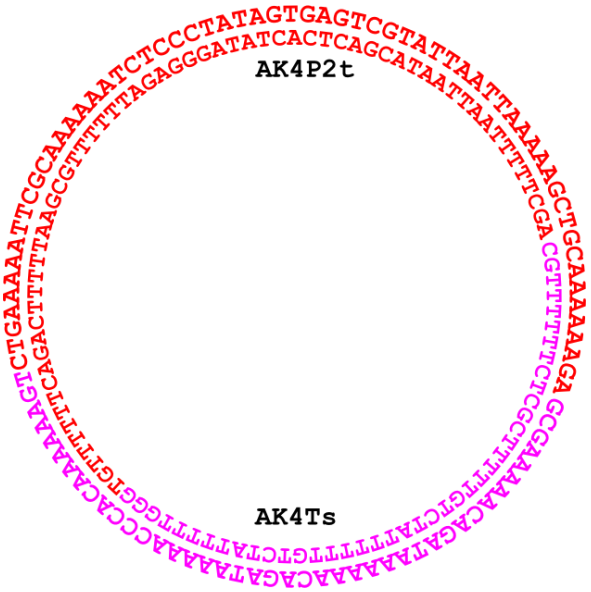
Promotor segment AK4Pt, removed fragment AK4Oa, AK4Ob, 126 bp

```
AK4T1
  CCCAAAAAGCCACAAAAAACGGCGAAAAACAGATAAAAAACAGATAAAAACCCACAAAAAAGT
CGTTTTTTCTGGGTTTTTCGGTGTTTTTTGCCGCTTTTTGTCTATTTTTTGTCTATTTTTGGG
     --------t --------t********
AK4P2t
  CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTGCAAAAAAGA
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGA
*******************************.....<--------------prrr
```

## AK45

Short termination segment AK4Ts, promotor segment AK4Pt, removed fragment AK4Oa, AK4Ob, 105 bp

```
AK4Ts
  GCGAAAAACAGATAAAAAACAGATAAAAACCCACAAAAAAGT
CGTTTTTTCTCGCTTTTTGTCTATTTTTTGTCTATTTTTGGG
  --------t --------t********
AK4P2t
  CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTGCAAAAAAGA
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGA
*******************************.....<--------------prrr
```



## Expected transcription products AK4 circles:

First termination site 40%

56n GGGAGATTTTTTGCGAATTTTTCAGACTTTTTTGTGGGTTTTTATCTGTTTTTTAT

Second termination site 24%

68n GGGAGATTTTTTGCGAATTTTTCAGACTTTTTTGTGGGTTTTTATCTGTTTTTTATCTGTTTTTCGCT

First run around termination at first site
161 - 266 n depending on circle size 14.4%

First run around termination at second site
173 - 278 n depending on circle size 8.64%

Second run around termination, both terminators
~350 - 450 depending on circle size 8.29%

Third run around transcription, both terminators
~430 - 630 depending on circle size 2.99%

Third run around transcription, both terminators
~610 - 910 depending on circle size 1.00%

## AK410b

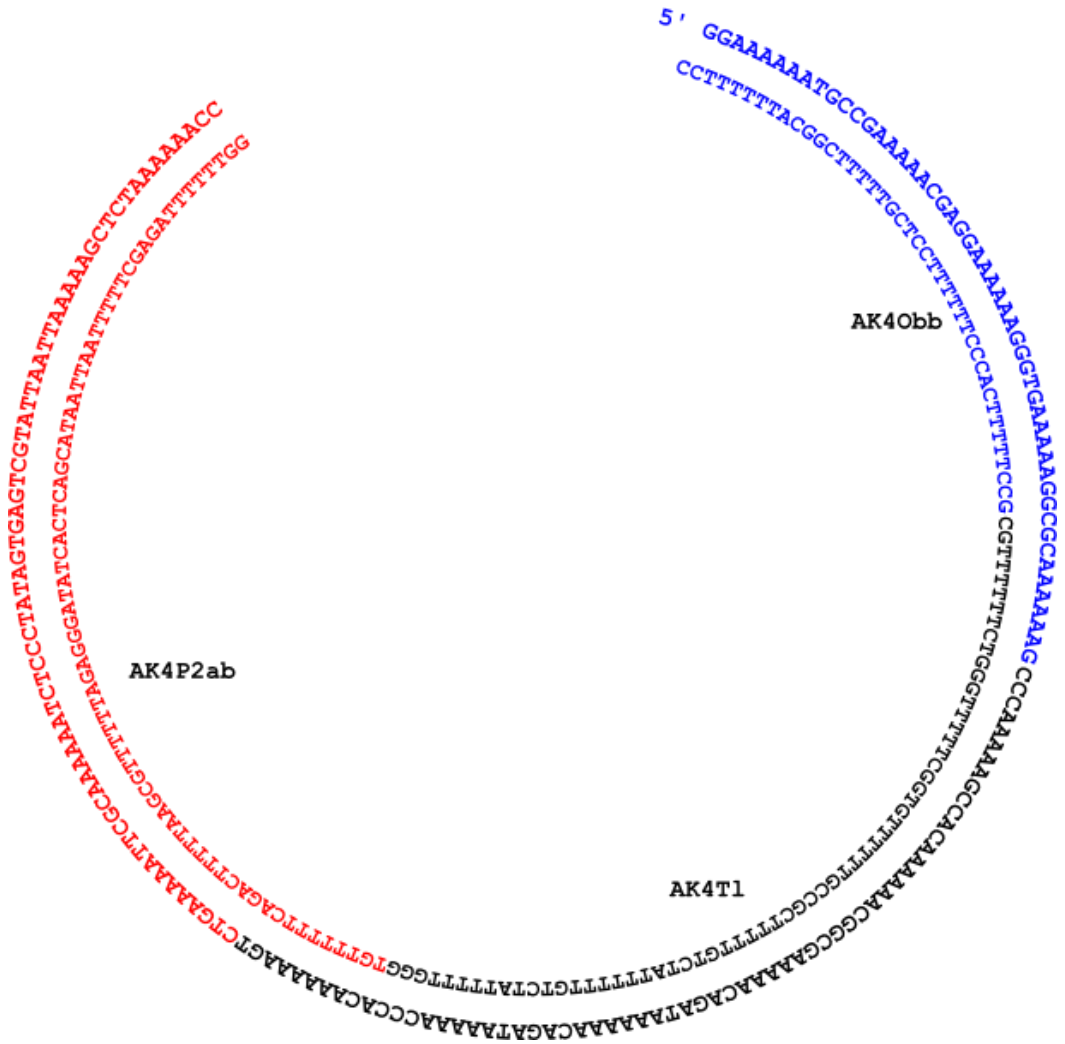
Blunt end fragment of circle AK410 omitting fragment AK4Oa, blunt segments AK4Pab and AK4Obb, 178 bp

```
AK4T1
  CCCAAAAAGCCACAAAAAACGGCGAAAAACAGATAAAAAACAGATAAAAACCCACAAAAAAGT
CGTTTTTTCTGGGTTTTTCGGTGTTTTTTGCCGCTTTTTGTCTATTTTTTGTCTATTTTTGGG
     --------t --------t********

AK4P2ab
  CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTCTAAAAAACC
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGAGATTTTTTGG
*******************************.....<--------------prrr

AK4Obb

GGAAAAAATGCCGAAAAACGAGGAAAAAAGGGTGAAAAAGGCGCAAAAAAGA
CCTTTTTTACGGCTTTTTGCTCCTTTTTTCCCACTTTTTCCG
```



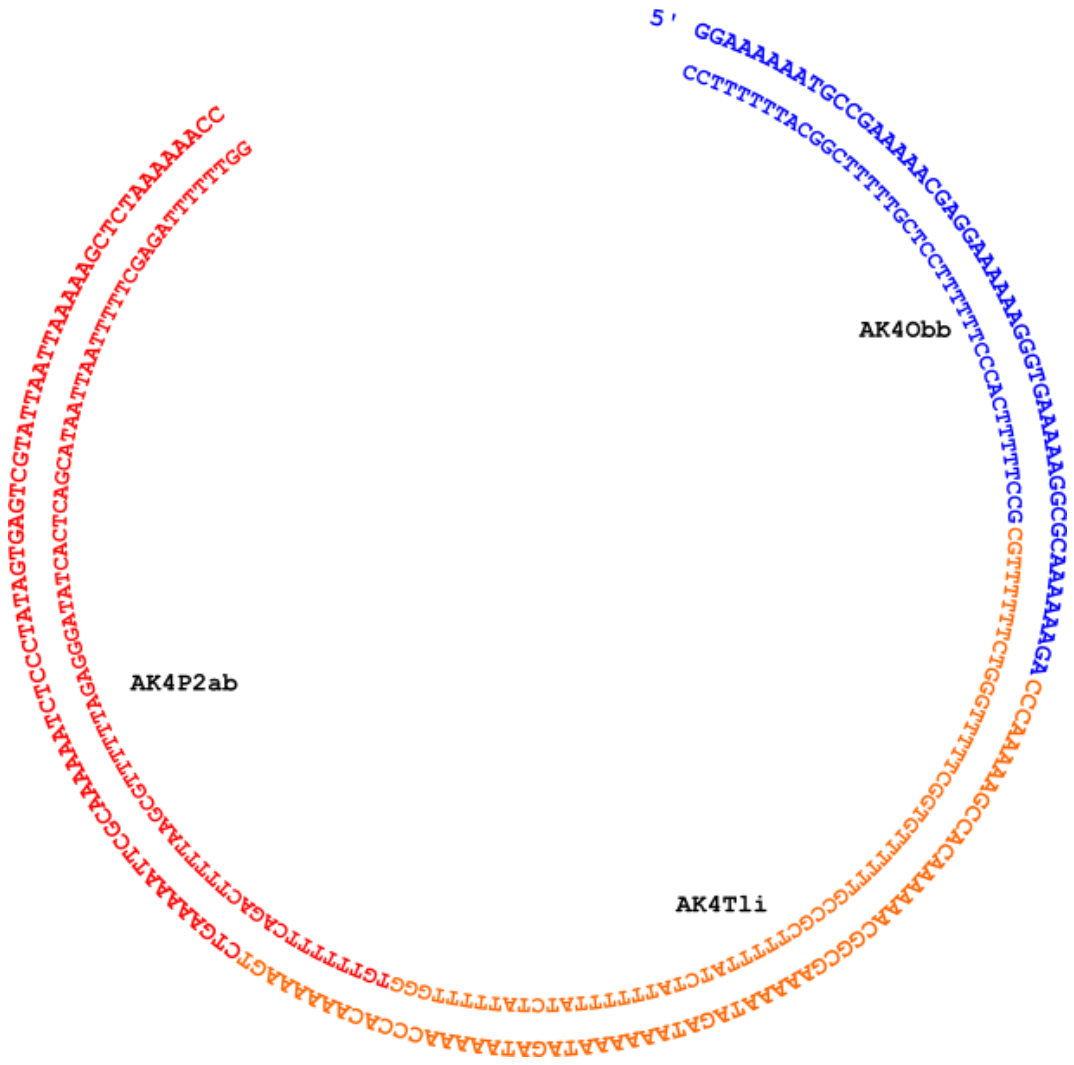
## AK410bi

Blunt end fragment of circle AK410 omitting fragment AK4Oa, blunt segments AK4Pab and AK4Obb, non functional terminator in fragment AK4Tli, 178 bp

```
Ak4Tli
  CCCAAAAAGCCACAAAAAACGGCGAAAAATAGATAAAAAATAGATAAAAACCCACAAAAAAGT
CGTTTTTTCTGGGTTTTTCGGTGTTTTTTGCCGCTTTTTATCTATTTTTTATCTATTTTTGGG
************************************************************

AK4P2ab
  CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTCTAAAAAACC
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGAGATTTTTTGG
*****************************.....<--------------prrr


AK4Obb
GGAAAAAATGCCGAAAAACGAGGAAAAAAGGGTGAAAAAGGCGCAAAAAAGA
CCTTTTTTACGGCTTTTTGCTCCTTTTTTCCCACTTTTTCCG
```



## Expected transcription products AK410b and AK410bi blunt end fragments

### *AK410b*

First termination site 40%

56n GGGAGATTTTTTGCGAATTTTTCAGACTTTTTTGTGGGTTTTTATCTGTTTTTTAT

Second termination site 24%

68n GGGAGATTTTTTGCGAATTTTTCAGACTTTTTTGTGGGTTTTTATCTGTTTTTTATCTGTTTTTCGCT

Run of transcription 36%

140n 5'
GGGAGATTTTTTGCGAATTTTTCAGACTTTTTTGTGGGTTTTTATCTGTTTTTTATCTGTTTTTCGCCGTTTTTTGT
GGCTTTTTGGGTCTTTTTTGCGCCTTTTTCACCCTTTTTTCCTCGTTTTTCGGCATTTTTTCC

### *AK410bi*

Run of transcription 100%

140n 5'
GGGAGATTTTTTGCGAATTTTTCAGACTTTTTTGTGGGTTTTTATCTGTTTTTTATCTGTTTTTCGCCGTTTTTTGT
GGCTTTTTGGGTCTTTTTTGCGCCTTTTTCACCCTTTTTTCCTCGTTTTTCGGCATTTTTTCC

## All fragments and oligos*:*

Oligos in 5′ → 3′ direction

### *AK4T*:

AK4Tl

CCCAAAAAGCCACAAAAAACGGCGAAAAACAGATAAAAAACAGATAAAAACCCACAAAAAAGT
CGTTTTTTCTGGGTTTTTCGGTGTTTTTTGCCGCTTTTTGTCTATTTTTTGTCTATTTTTGGG

AK4Ts

GCGAAAAACAGATAAAAAACAGATAAAAACCCACAAAAAAGT
CGTTTTTTCTCGCTTTTTGTCTATTTTTTGTCTATTTTTGGG

Ak4Tli

CCCAAAAAGCCACAAAAAACGGCGAAAAATAGATAAAAAATAGATAAAAACCCACAAAAAAGT

CGTTTTTTCTGGGTTTTTCGGTGTTTTTTGCCGCTTTTTATCTATTTTTTATCTATTTTTGGG

Oligos AK4T:

AK4Tl_f: CCCAAAAAGCCACAAAAAACGGCGAAAAACAGATAAAAAACAGATAAAAACCCACAAAAAAGT

Ak4Tl_r: GGGTTTTTATCTGTTTTTTATCTGTTTTTCGCCGTTTTTTGTGGCTTTTTGGGTCTTTTTTGC

Ak4Ts_f: GCGAAAAACAGATAAAAAACAGATAAAAACCCACAAAAAAGT

Ak4Ts_r: GGGTTTTTATCTGTTTTTTATCTGTTTTTCGCTCTTTTTTGC

Ak4Tli_f: CCCAAAAAGCCACAAAAAACGGCGAAAAATAGATAAAAAATAGATAAAAACCCACAAAAAAGT

Ak4Tli_r: GGGTTTTTATCTATTTTTTATCTATTTTTCGCCGTTTTTTGTGGCTTTTTGGGTCTTTTTTGC

## *AK4P:*

AK4P2a

CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTCTAAAAAACC
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGA

AK4P2b

CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTGGAAAAAATG
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGA

AK4P2t

CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTGCAAAAAAGA
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGA

AK4P2ab

CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTCTAAAAAACC
TGTTTTTTCAGACTTTTTAAGCGTTTTTTAGAGGGATATCACTCAGCATAATTAATTTTTCGAGATTTTTTGG

## *Oligos AK4P2*

AK4P2a_f: CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTCTAAAAAACC

AK4P2_r: AGCTTTTTAATTAATACGACTCACTATAGGGAGATTTTTTGCGAATTTTTCAGACTTTTTTGT

AK4P2b_f: CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTGGAAAAAATG

AK4P2t_f: CTGAAAAATTCGCAAAAAATCTCCCTATAGTGAGTCGTATTAATTAAAAAGCTGCAAAAAAGA

AK4P2ab_r:
GGTTTTTTAGAGCTTTTTAATTAATACGACTCACTATAGGGAGATTTTTTGCGAATTTTTCAGACTTTTTT
GT

### *AK4Oa*

ACCAAAAAGGAGGAAAAAACGTGGAAAAATGCGGAAAAAATG
GATTTTTTGGTGGTTTTTCCTCCTTTTTTGCACCTTTTTACG

### *Oligos AK4Oa*

AK4Oa_f: ACCAAAAAGGAGGAAAAAACGTGGAAAAATGCGGAAAAAATG

AK4Oa_r: GCATTTTTCCACGTTTTTTCCTCCTTTTTGGTGGTTTTTTAG

AK4Ob

AK4Ob

CCGAAAAACGAGGAAAAAAGGGTGAAAAAGGCGCAAAAAAGA
CCTTTTTTACGGCTTTTTGCTCCTTTTTTCCCACTTTTTCCG

AK4Obb

GGAAAAAATGCCGAAAAACGAGGAAAAAAGGGTGAAAAAGGCGCAAAAAAGA
CCTTTTTTACGGCTTTTTGCTCCTTTTTTCCCACTTTTTCCG

Oligos AK4Ob/bb

AK4Ob_f: CCGAAAAACGAGGAAAAAAGGGTGAAAAAGGCGCAAAAAAGA

Ak4Ob_r: GCCTTTTTCACCCTTTTTTCCTCGTTTTTCGGCATTTTTTCC

AkfObb_f: GGAAAAAATGCCGAAAAACGAGGAAAAAAGGGTGAAAAAGGCGCAAAAAAGA

## Appendix 3: Point kinetics

Linear regression for 100 nM of circular templates 126 to 210 bp in circumference. See: Materials and methods: Real time point assay of template concentrations of 100 nM.