Appl Biochem Biotechnol (2011) 164:497–513
DOI 10.1007/s12010-010-9152-4

Early In Vitro Transcription Termination in Human H5 Influenza Viral RNA Synthesis

Matthew B. Kerby · Aartik A. Sarma · Madhukar S. Patel · Andrew W. Artenstein · Steven M. Opal · Anubhav Tripathi

Received: 28 July 2010 / Accepted: 20 December 2010 /
Published online: 5 January 2011

Abstract Rapid diagnostic identification of the human H5 influenza virus is a strategic cornerstone for outbreak prevention. We recently reported a method for direct detection of viral RNA from a highly pathogenic human H5 influenza strain (A/Hanoi/30408/2005 (H5N1)), which necessarily was transcribed in vitro from non-viral sources. This article provides an in-depth analysis of the reaction conditions for in vitro transcription (IVT) of full-length influenza H5 RNA, which is needed for diagnostic RNA production, for the T7 and SP6 phage promoter systems. Gel analysis of RNA transcribed from plasmids containing the H5 sequence between a 5′ SP6 promoter and 3′ restriction site (BsmBI) showed that three sequence-verified bands at 1,776, 784, and 591 bases were consistently produced, whereas only one 1,776-base band was expected. These fragments were not observed in H1 or H3 influenza RNA transcribed under similar conditions. A reverse complement of the sequence produced only a single band at 1,776 bases, which suggested

M. B. Kerby · A. A. Sarma · M. S. Patel · A. Tripathi
School of Engineering and Division of Biology and Medicine, Biomedical Engineering, Center for Biomedical Engineering, Brown University, Providence, RI 02912, USA

A. W. Artenstein · S. M. Opal · A. Tripathi
Department of Medicine & Center for Biodefense and Emerging Pathogens, Memorial Hospital of RI, Pawtucket, RI, USA

A. W. Artenstein · S. M. Opal · A. Tripathi (✉)
Warren Alpert School of Medicine, Brown University, Box D, Providence, RI 02912, USA
e-mail: anubhav_tripathi@brown.edu

Present Address:
M. B. Kerby
Department of Bioengineering, Stanford University, Stanford, CA, USA

Present Address:
A. A. Sarma
Harvard Medical School, Boston, MA, USA

Present Address:
M. S. Patel
School of Medicine, University of California, Irvine, Irvine, CA, USA

either self-cleavage or early termination. Aliquots of the IVT reaction were quenched with EDTA to track the generation of the bands over time, which maintained a constant concentration ratio. The H5 sequence was cloned with T7 and SP6 RNA polymerase promoters to allow transcription in either direction with either polymerase. The T7 transcription product from purified, restricted plasmids in the vRNA direction only produced the 1,776-base full-length sequence and the 784-base fragment, instead of the three bands generated by the SP6 system, suggesting an early termination mechanism. Additionally, the T7 system produced a higher fraction of full-length vRNA transcripts than the SP6 system did under similar reaction conditions. By sequencing we identified a type II RNA hairpin loop terminator, which forms in a transcription direction-dependent fashion. Variation of the magnesium concentration produced the greatest impact on termination profiles, where some reaction mixtures were unable to produce full-length transcripts. Optimized conditions are presented for the T7 and SP6 phage polymerase systems to minimize these early termination events during in vitro transcription of H5 influenza vRNA.

Keywords Termination · Influenza · H5 · In vitro transcription · SP6 · T7

Introduction

Continuing outbreaks and evolution of avian influenza virus reinforce the need for a field portable diagnostic device to provide information on viral gene sequences [1], which can provide information about pathogenicity and drug resistance to help mediate a proper health response. PCR, the de facto standard for amplification of genetic material, requires thermocycling with post-amplification electrophoresis or assays using specialized lab equipment for real-time RT-PCR. While portable PCR devices are under development [2], established non-PCR-based assays, such as nucleic acid sequence-based amplification (NASBA) [3] or antisense RNA microarray diagnostic methods [4], rely on RNA transcripts between primers to function. Since template sequences and reaction conditions modulate transcription pausing and early termination [5–7], these diagnostic test methods often lose effectiveness or fail without extensive optimization. Thus, diagnostic RNA chemistries including reverse transcription for viral genes, which have extensive secondary structures and/or high GC content [8, 9], are especially challenging. These systems often use the T7 and SP6 RNA polymerases, which show high specificity for their respective promoters and good processivity [10], which has led to their extensive use for bacterial protein expression [11–13], in vitro transcription (IVT) reactions [14–18], and incorporation into diagnostic assays [19–22]. We sought to use these polymerases to generate vRNA sequences for evaluation of systems to detect various subtypes of influenza and were able to amplify full-length H1, H3, N1, and N2 genes. Two shorter fragments of RNA were observed in systems meant to amplify full-length H5 influenza RNA (NCBI-AB239125), and this article investigates the optimal method to synthesize full-length sequences and the mechanisms underlying the creation of shorter sequences.

RNA is isothermally amplified by T7 and SP6 DNA-dependent RNA polymerases, which transcribe RNA from DNA templates. The amplified RNA may be diagnostically identified using probe arrays [23] or molecular beacons [24] for identification of viral subtype, hemagglutinin (HA) cleavage site sequence, and receptor specificity, which are all related to viral pathogenicity. The T7 and SP6 phage promoters begin transcription on double-stranded DNA templates, then displace the top DNA strand and use the 3′ strand as a template for RNA production. Transcription can be slowed by one of two regulatory

pause mechanisms: RNA hairpins or backtracking. Backtracking pauses occur with weak RNA/DNA sequence hybrids, which cause the transcription complex to move in reverse and force the nascent RNA transcript back through the channel where nucleotides are added [25]. Hairpin pauses are induced by self-complementary transcripts, which inhibit forward motion of the transcription complex and prevent the addition of new nucleotides [26].

In addition to regulator pauses, the T7 and SP6 RNA polymerases also respond to two classes of terminator sequences that trigger transcript release. Termination signals are characterized as either RNA-dependent hairpin structures or DNA-dependent sequences [27]. These rho-independent intrinsic terminators of *Escherichia coli* RNA polymerase (RNAP) occur between genes 10 and 11 in the T7 genome and appear to be important in attenuating transcription of downstream genes [28, 29]. Evidence suggests that the hairpin formation at these terminators can disrupt RNAP interactions with the single-stranded segment of RNA 8–14 nucleotides (nt) away from the RNA 3′ end [27]. The hairpin formation can also disrupt some of the base pairing in the RNA/DNA hybrid which results in a weak residual U-rich hybrid unable to hold the RNA in the transcription complex [30]. This type-1 termination caused by RNA secondary structure is reduced significantly or eliminated by substituting guanine triphosphate with the nucleotide inosine triphosphate, where a non-hydrogen pairing H replaces an NH_3 group [31]. Dobbins et al. [32] provide the locations for only seven putative termination sequences in the SP6 genome (AY288927) they sequenced, but none of these strong hairpin RNA sequences is present in the H5 influenza gene of NCBI-AB239125.

The phage polymerases have also been shown to terminate transcription in response to features of the DNA template. One sequence, ATCTGTT, has been shown to act as a pause site and has been shown to cause termination if it is followed by a T-rich sequence [33–36]. Portugal and Rodriguez-Campos [37] analyzed the ability of T7 RNA polymerase to transcribe a plasmid DNA in vitro in linear, supercoiled, relaxed, and knotted DNA forms; knotting was the only topology to significantly impact transcription of RNA.

Having discussed regulatory pausing and termination signals, it is important to consider the self-cleaving activity of RNA that is exemplified by the hammerhead ribozyme, which can facilitate the hydrolysis of other RNA molecules. In addition, cleavage activity of human influenza RNA has been reported in vivo for viral replication where the polyadenylated tails of host messenger RNA are cleaved and appended to the viral RNA (vRNA) segments for host translation [38]. This process, known as cap-snatching, involves the influenza RNA-dependent RNA polymerase.

Commercial in vitro RNA synthesis chemistries are optimized for a standard set of conditions using model templates but can produce large numbers of unexpected products with different templates. This investigation seeks to determine the cause of the shortened viral RNA fragments and provide a solution for synthesis of full-length vRNA for diagnostic assays for H5 influenza. The transcription product ratio over time was observed while evaluating the effects of the H5 influenza template sequence, phage promoter type, and solution chemistry. RNA synthesis disruption by an RNA cleavage mechanism and early termination of RNA synthesis were explored.

Methods

Preparation of RNA Templates

A wild-type human-isolate influenza A H5 DNA sequence (A/Hanoi/30408/2005 (H5N1) NCBI-AB239125) was synthesized (DNA 2.0, Menlo Park, CA, USA), inserted into a pJ10

E. coli propagation plasmid, and sequenced. The RNA of the influenza virus encoding the HA segment was 1,776 bases with two highly conserved ends. A truncated SP6 RNA polymerase promoter, which lacks a G at position 1, was followed by the reverse complement of the H5 cDNA sequence. The 3′ end of the sequence includes a restriction site for BspEI/Kpn2I, which cleaves the sequence "TCCGGA" one base in from the 5′ ends on both strands. A second cleavage site was included upstream of the SP6 promoter site. Plasmid cleavage with this enzyme produces a 3′ overhang that was removed using the Klenow fragment of DNA polymerase I. The 5′ cleavage site was simultaneously filled in to produce a template for H5 viral RNA with native 5′ and 3′ ends. Transcription products, produced using an SP6 Megascript kit (Ambion, Austin, TX, USA), were visualized on an RNA gel microchip (Agilent, RNA 6000 Nano). To ensure consistency among the experiments with RNA transcripts, a set of four complementary clones, two with SP6 and two with T7 promoters, were created and contained the H5 influenza HA gene with matching promoters as shown in Fig. 1.

These constructs produced RNA products corresponding to the native ends found with the influenza virus plus a single G residue at position 1. Use of G as the first translated base has been shown to greatly increase the transcription efficiency. The primers, which contain restriction sites and T7 or SP6 promoter sites, are shown in Table 1. In the infectious viral state, influenza RNA has a highly conserved sequence beginning with AGU (AGT for DNA) that was used here for naming convention to identify clones that produce vRNA transcripts. Complementary RNA (cRNA), which is produced by the virus during the replication phase, has a sequence starting with AGC. This code will designate clones for cRNA production, which are the reverse complement to vRNA clones. The clone primer pairs used from Table 1 are shown in parenthesis following each clone name: AGT-T7 (3,4), AGC-T7 (6,1), AGT-SP6 (2,4), and AGC-SP6 (5,1). Sequences were created from the synthesized H5 plasmid and amplified via PCR using Phusion high-fidelity DNA polymerase (New England Biolabs, Ipswich, MA, USA). PCR was conducted in a BioRad MyCycler (Hercules, CA, USA) at maximum ramp rates with initial denature of 98 °C for 30 s, then 30 cycles of 98 °C for 10 s, anneal for 15 s at T_m +3 °C of the lowest primer for the combination, and 45 s at 72 C.

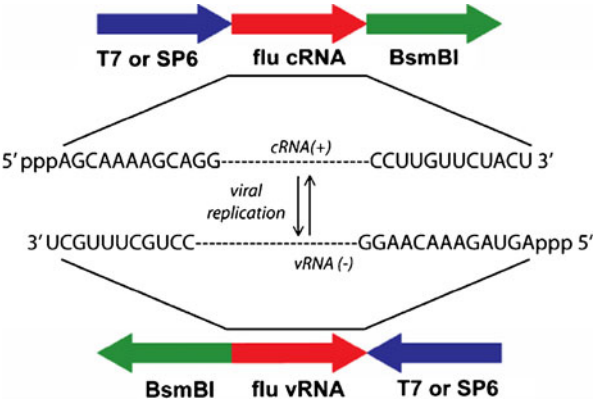

Fig. 1 H5 clone inserts allow synthesis of H5 in two directions via SP6 or T7 chemistries. The *arrows* represent gene segments assembled to create H5 clones. In vivo, influenza viral cRNA and vRNA maintain highly conserved ends for replication. Clones designated AGT produce vRNA and AGC clones produce cRNA. BsmBI is a restriction site positioned to cut the template strand to create a native-end RNA

Table 1 Primer list for H5 clone creation

	Primer	Sequence	T_m(°C)
1	AGT_BtgZI-BsmBI-HA	GCGATGAACCGTCTCT **AGTAGAAACAAGGGTGTTTT**	63.3
2	AGT_**SP6**-G-HA	ATTTAGGTGACACTATAG **AGTAGAAACAAGGGTGTTTT**	60.1
3	AGT_**T7**-G-HA	CGGCTAATACGACTCACTATAG **AGTAGAAACAAGGGTGTTTT**	62.9
4	AGC_BtgZI-BsmBI-HA	GCGATGAACCGTCTCT **AGCAAAAGCAGGGG**	65.4
5	AGC_**SP6**-G-HA	CGGCATTTAGGTGACACTATAG **AGCAAAAGCAGGGG**	65.0
6	AGC_**T7**-G-HA	TAATACGACTCACTATAG **AGCAAAAGCAGGGG**	60.5

The sequences in bold are conserved in influenza

The amplified sequences were cloned into a pJet1 plasmid (Fermentas) and transfected into New England Biolabs (NEB) 10-beta *E. coli*. Colonies were screened using orientational PCR such that the phage promoter site was adjacent to the Kpn2I restriction site of the plasmid. The restriction sites for runoff transcription are recognized by Type IIs restriction enzymes, which recognize non-palindromic sequences and cleave outside of the recognition site. The 3′ primer contained a BsmBI/BtgZI site, which provided flexibility in cleavage enzyme selection but cut the same location. Single PCR-positive colonies of *E. coli* were shaken overnight at 37 °C in 5 mL of Luria broth with 100 μg/mL of carbenicillin. Master stocks of active bacteria were stored at −80 °C in 15% glycerol. Plasmids were harvested using the Wizard Plus SV Minipreps DNA purification system (Promega) and sequenced in both directions to validate the construct. The plasmids were linearized with BsmBI (New England Biolabs), a *dam* methylation insensitive restriction enzyme, for 1 h at 55 C and purified using the Wizard SV gel and PCR Cleanup kit (Promega). Yield was quantified spectrophotometrically on a Nanodrop 1000 (Nanodrop, Wilmington, DE, USA). Materials were stored at −20 °C until needed. These plasmid-derived DNA sequences, shown in Fig. 1, were used for generating non-pathogenic viral RNA molecules for this investigation.

DNA and RNA Analysis

Quantitative analysis for both DNA and RNA was completed using an Agilent 2100 Bioanalyzer running the DNA 7500, RNA 600 Nano, or RNA 6000 Pico labchip kits. Reagent concentrations were verified to fall within the range specified for quantitative analysis. Secondary structure of RNA was calculated using MFOLD server [39] assuming 6 mM [Mg^{2+}] to locate regions of the strongest base pairing interactions.

In Vitro Transcription

Runoff transcripts of H5 influenza RNA were produced with 500 ng of purified DNA template in 20 μL volumes. Preformulated Megascript kits (Ambion) for SP6 and T7 were used at 7.5 mM total rNTP as per direction, but the [Mg^{2+}] was withheld. The highly tunable, multicomponent T7 and SP6 systems from NEB were used in parallel. The NEB reaction mixture was assembled according to the manufacturer's recommendation and contained 40 mM Tris–HCl, 6 mM $MgCl_2$, 10 mM dithiothreitol, 2 mM spermidine, pH 7.9 at 25 °C, 2 mM total rNTP, 1 U/20 μL reaction inorganic pyrophosphatase, 100 μg/mL purified bovine serum albumin, 1 U/20 μL reaction of RNAse inhibitor, and 1 U/20 μL

reaction of polymerase. Additional tests of elevated [Mg^{2+}], [rNTP], or temperature were performed using the NEB chemistry. According to NEB, the SP6 RNA polymerase is 30% more active at 40 °C than at 37 °C. However, all reactions were incubated at 37 °C unless otherwise specified.

Time Series In Vitro Transcription Reactions

IVT reactions in SP6 and T7 were repeated using the four clones for H5 vRNA and cRNA, where samples were withdrawn at fixed time intervals and quenched in 10 mM EDTA to chelate the magnesium and stop both transcription and possible self-cleavage activity. Following the collection of all samples, the DNA template was degraded with 1 U/10 μL DNAse in an equal volume of 15 mM $MgCl_2$ and incubated at 37 °C for 20 min. Samples were chilled on ice and immediately sized and quantified on an Agilent RNA 6000 Nano chip. The time corrected area for each minor peak of the electropherogram was divided by the time corrected area of the major expected peak.

Reverse Transcription

The short transcription products were by sequencing reverse-transcribed DNA products created using a modified RACE-PCR protocol. The RNA products were polyadenylated with *E. coli* poly(A) polymerase (NEB) following the manufacturer's instructions. The sequence was reverse-transcribed using Thermoscript Reverse Transcriptase (Invitrogen) using a high T_m primer to bind to the 5′ end of the poly-A extension (GCG AGC ACA GAA TTA ATA CGA CTC ACT ATA $GG(T)_{18}$ VN). After treatment with RNase H (Invitrogen), the reverse-transcribed DNA was amplified by PCR using the above primer and a primer for the 5′ end of the AB239125 sequence (AGT AGA AAC AAG GGT GTT TTT AAC TAC AAT CTG). The products were purified and separated using a 2.0% agarose gel, and the excised gel fragments were purified using a Wizard SV PCR cleanup kit. The isolated fragments were then amplified again to ensure purity and produce sufficient sample for sequencing. The DNA products were sequenced by the WM Keck Foundation Biotechnology Resource Laboratory at Yale University. The chromatograms were aligned with the known sequence on CLC Main Workbench 4.1.1 (CLC Bio USA, Cambridge, MA, USA), and the 3′ sequences of the short fragments were determined.

Results

Runoff transcripts of the restriction digest of the pJ10 AGT-SP6-H5 plasmid consistently resulted in the generation of three bands. To isolate the source of these bands, the template, the sequence orientation, the phage promoter, and the solution chemistry were investigated.

Plasmid Design

For PCR reactions using the Phusion DNA polymerase, an annealing temperature 3°C above the T_m of the primer reaction drastically reduced the formation of primer–dimers in the product formed by the palindromic "CTATAG" sequence present in both the T7 and SP6 primers. Taq chemistries, which operated at lower annealing temperatures, produced

poor yields and required gel purification. The restriction digest was separated on an Agilent 7500 DNA chip and showed, as expected, two fragments near 1,776 and 2,241 bp.

Template Restriction or Multiple Promoter Sites Are Not Responsible for Multiple RNA Bands

RNA transcription of native-end H5 influenza from the restricted pJ10 AGT-SP6-H5 plasmid resulted in three band fragments as shown in Fig. 2. Post-restriction digest DNA gels had two clean bands of expected sizes; interestingly, however, three RNA bands were still produced following a 4-h SP6 IVT reaction at 37 °C. Template preparation was not a likely cause of the vRNA banding since both the methylation-sensitive BspEI (New England Biolabs) and the non-methylation-sensitive Kpn2I (Fermentas) restriction enzymes produced identical DNA plasmid fragments. Next, the plasmid was checked for other sites capable of priming an SP6 transcription. The pJ10 plasmid has one site with a 7-base sequence match to the SP6 promoter but does not contain the full sequence. To eliminate any possible plasmid promoter contribution, an AGT-SP6-H5 segment was amplified using the Phusion polymerase (Finnzymes), which has 3′ exonuclease activity, to produce a high purity template verified by a single band on a DNA gel shown in Fig. 2a. Following Megascript SP6 transcription, the RNA product still contained three bands shown in Fig. 2b. The H5 RNA IVT solution was reverse-transcribed to first-strand cDNA and amplified via PCR to confirm full-length RNA was present as shown in Fig. 2c.

RNA Band Ratio Remained Constant in Time

In vitro transcription was observed over time for a 16-h reaction at 37 °C. A time series of SP6 IVT was performed using Kpn2I-restricted plasmid and PCR amplified template. Aliquots were withdrawn at 1, 2, 4, 6, 8, and 16 h time points from a single reaction mixture and gel analyzed. Figure 3 clearly shows that the three bands are present at all recorded time points, independent of the DNA template construction method. The results

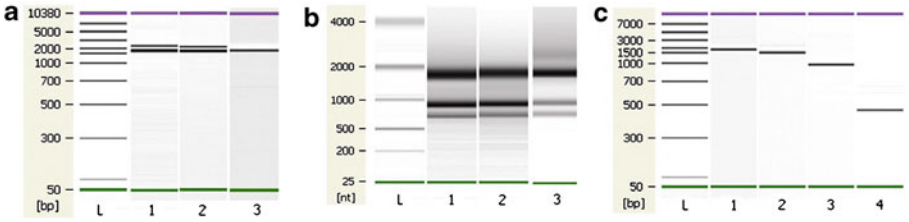

Fig. 2 Early termination of H5 vRNA synthesis is independent of restriction enzyme and plasmid template. Transcription from multiple templates using a Megascript SP6 kit. **a** A composite DNA pseudogel from an Agilent 2100 of restricted H5 plasmids with sizes as expected: *lane 1* BspEI digest, *lane 2* Kpn2I digest, *lane 3* SP6-H5 PCR fragment exactly matching the cleaved insert. The templates are used for production of influenza H5 RNA via SP6 in vitro transcription. **b** A composite RNA pseudogel of H5 RNA produced by SP6 in vitro transcription of DNA templates. *Lane 1* BspEI digest, *lane 2* Kpn2I digest, *lane 3* H5 PCR product. RNA products shown at 1,776 (full-length H5 vRNA), 784, and 591. **c** Full-length H5 is present. H5 IVT vRNA was reverse-transcribed to cDNA using a consensus H5 primer designed to produce the full-length transcript of 1,776 nt, then amplified via PCR using the consensus 5′ H5 primer and 3′ primers that produce (*lane 1*) 1,776, (*lane 2*) 1,475, (*lane 3*) 964, and (*lane 4*) 464 nt products

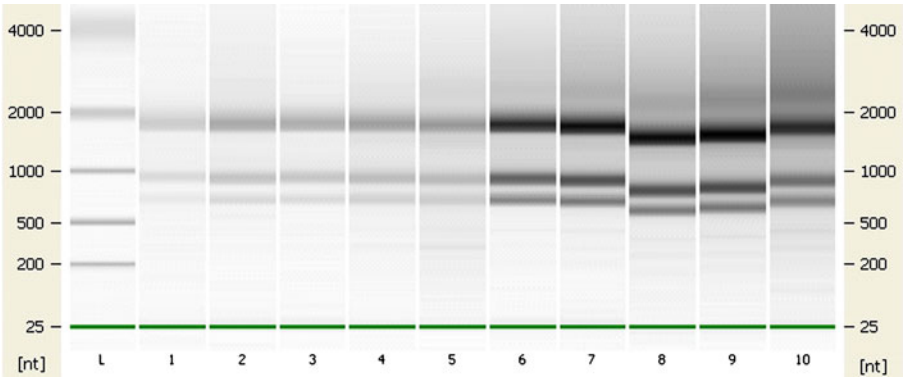

Fig. 3 Time series of RNA production. Samples from two IVT reactions using templates from Kpn2I-restricted cDNA (*lanes 1–5*) or H5 PCR product (*lanes 6–10*) were withdrawn at 1, 2, 4, 6, and 16 h intervals. The bands are present in both systems, which eliminate the possibilities of incomplete restriction digest or contaminating DNA products. The band shift in *lanes 8* and *9* is an artifact of the software analysis of those lanes and not a real difference

suggest incomplete restriction digest or contaminating DNA products are not factors in production of the bands.

Creation of Four Symmetric H5 Clones with Matching Ends

The four H5 clones created from primers in Table 1 and inserted into matching pJet1 plasmids, not pJ10, were digested with BsmBI as shown in Fig. 4. These DNA bands match simulated digests and correlate with plasmid sequencing results. These clones equalize the effects of template across all constructs through identical starting and ending sequences.

H5 Banding Dependent on Sequence and Phage Promoter

Generation of multiple band fragments depended on the orientation of the H5 sequence when using the Megascript chemistry. The AGC clones, which synthesize cRNA, were-transcribed in parallel with the AGT clones that produce vRNA. Figure 5 shows electropherograms for the transcription products of all four clones after 4 h. Both the AGC-SP6 and AGC-T7 clones produce a single length of RNA, which appear as single

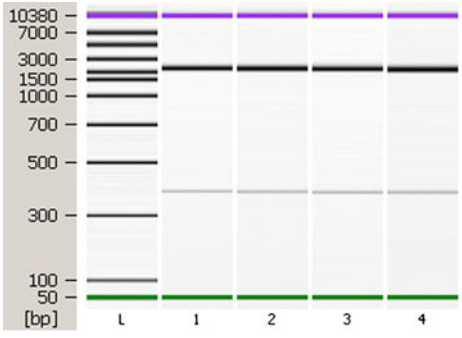

Fig. 4 A DNA pseudogel of plasmid DNA from four clones following a restriction digest by BsmBI. *Lane 1* AGT-T7-H5 vRNA template, *lane 2* AGC-T7-H5 cRNA template, *lane 3* AGT-SP6-H5 vRNA template, *lane 4* AGC-SP6 cRNA template

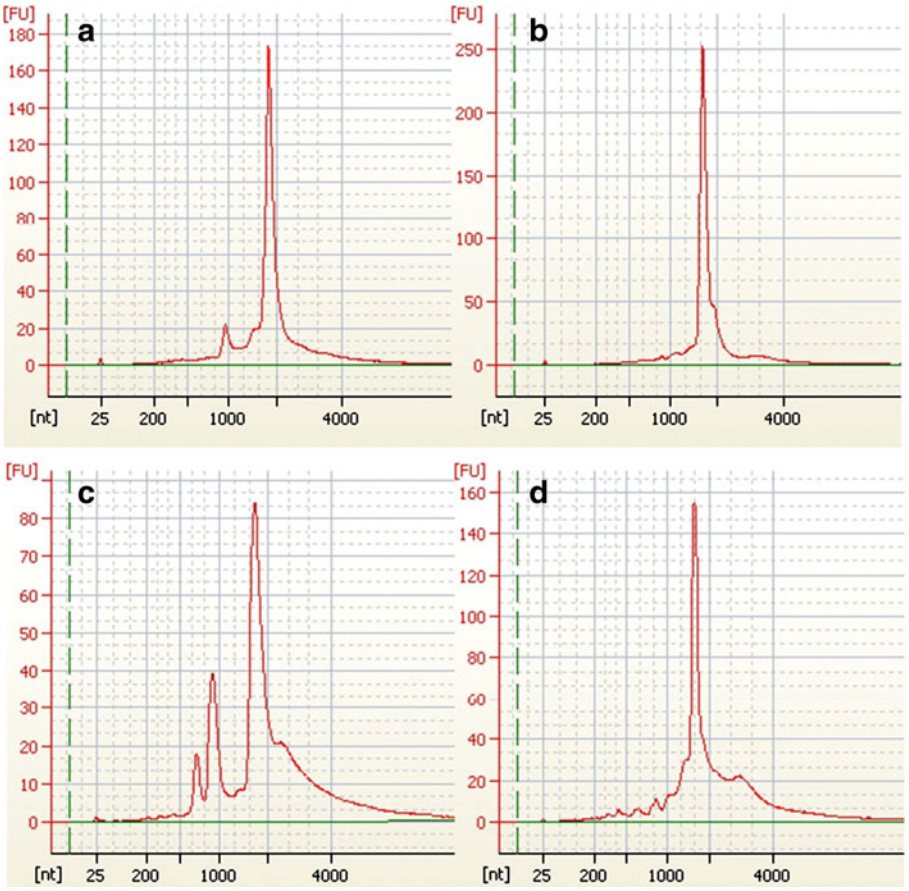

Fig. 5 Electropherograms of H5 RNA. **a** AGT-T7 vRNA, **b** AGC-T7 cRNA, **c** AGT-SP6 vRNA, **d** AGC-SP6 cRNA. RNA samples generated using the Megascript T7 and SP6 kits. Samples were pre-digested with DNAse prior to analysis. Each peak corresponds to the location of a band in a conventional gel view

peaks. Interestingly, the AGT-T7 clone produces two peaks at 784 and 1,776 bases whereas the AGT-SP6 clone again shows three peaks 1,776, 784, and 591 bases on the log-scale plot.

Reverse Transcription and Sequencing of RNA Products

Attempts using AMV reverse transcriptase to produce full-length RNA produced mixed fragments or no fragments, which was likely caused by the complex secondary structure of the H5 RNA sequence. The elevated temperature of the 60 °C RT-thermoscript kit (Invitrogen) improved first-strand transcription yields of influenza strands over 37 °C RT processing methods with AMV or MMLV. Based on the sequencing data, both SP6 and T7 polymerase produce a fragment ending at A784, and SP6 polymerase also produces a fragment ending at A591 (Fig. 6). Analysis of folded sequences processed by mFOLD suggested large hairpin structures and U-rich regions, although not sequence segment with highest U concentration, existed at both termination sites.

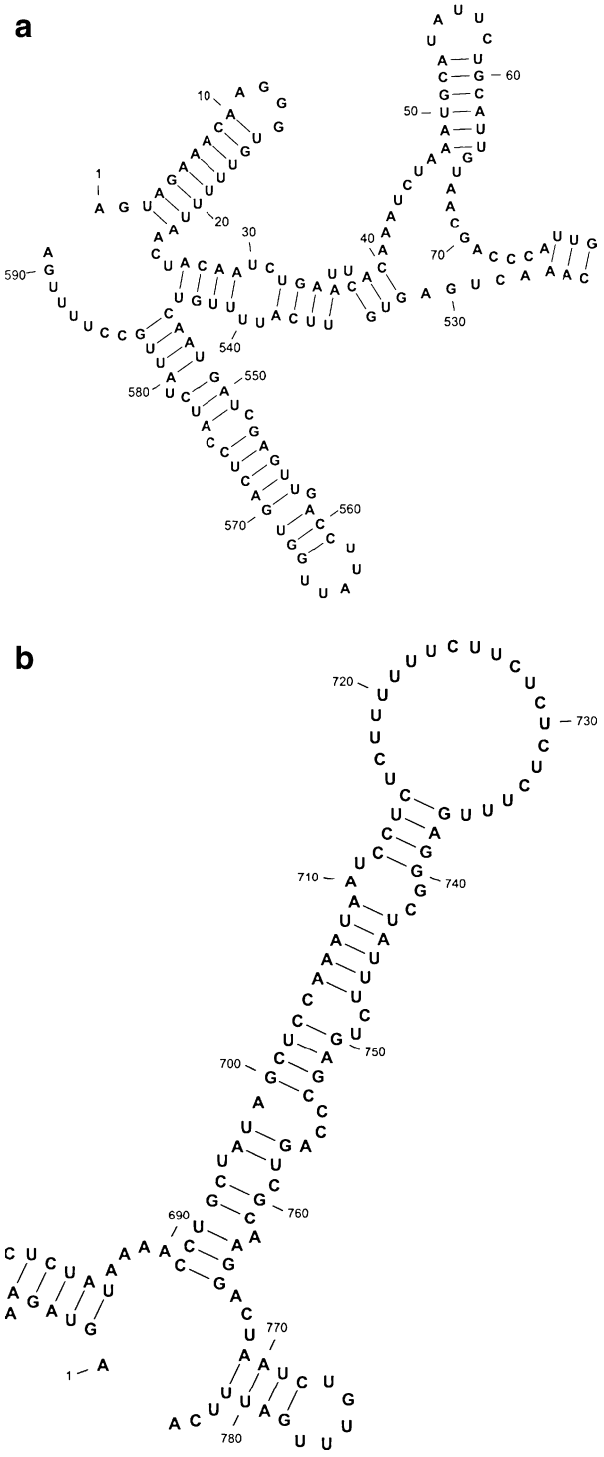

Fig. 6 RNA hairpin structures causing termination of transcription modeled by mFOLD. Sequences were determined using DNA synthesized by a modified rapid amplification of cDNA ends protocol. **a** Hairpin (bases 545–583) in 591-nt fragment produced by SP6 RNA polymerase. **b** Hairpin (bases 691–765) in 784-nt fragment produced by T7 and SP6 RNA polymerases

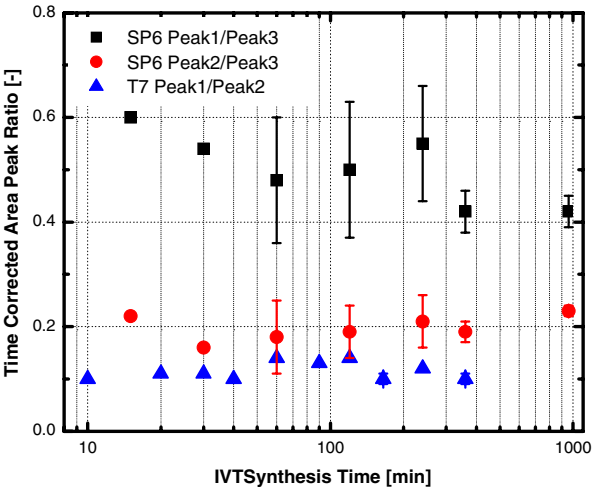

Fig. 7 Plot of the ratio of RNA fragments for the AGT-SP6-H5 and AGT-T7-H5 transcription reactions with Megascript chemistry. The ratio is computed from electropherogram data of time corrected area of the main product peak at 1,776 bases. The AGC-H5 fragments produced only a single peak and so are not shown. SP6 peak 1 is a 591-base transcript, SP6 peak 2 is a 784-base transcript, and SP6 peak 3 is the full-length 1,776-base transcript. T7 peak 1 is 784 bases and full length is 1,776 bases. *Error bars* represent standard deviation of multiple runs. $n=3$ at $t=60$, 120, and 240 min and $n=2$ at 360 and 960 min

H5 Banding Is Not Time Dependent

A time series of samples from Megascript transcription reactions were plotted as ratios of the peak area of the full-length transcript to the peak area of the shorter transcript. The ratio plot in Fig. 7 normalized the transcription rates for all constructs and revealed the changes in RNA accumulation at each transcript length. As a ratio of the SP6 transcription full-length RNA band, the 591-base transcript had a time corrected peak area 20±3% of the full-length transcript while the 784-base fragment averaged 50±6% of the area. The T7 transcription RNA peak area ratio averaged 12% of the full-length transcript. In both transcription systems, the fraction of by-products remained relatively constant over the transcription reaction, suggesting that cleavage was not occurring.

Incubation Temperature Alters the Termination Ratio

The AGT-SP6-H5 template was incubated at six temperatures of 25 °C, 30 °C, 32 °C, 36 °C, 37 °C, and 40 °C for 16 h. The gel and peak ratio for each DNAse digested reaction was plotted for each temperature in Fig. 8. Incubation at the coolest temperature of 25 °C reduced the area ratio of peak1/peak3 and peak2/peak3 by 85% and 59%, respectively. Transcription at 40 °C temperature greatly increased the fraction of short transcripts, which overtook the full-length transcript as the major transcription product.

Optimum Magnesium Concentration Reduces the Relative Concentration of Transcript Aborts

Magnesium concentration was increased from 6 to 16 mM total concentration in a 2-mM total rNTP mixture. The low magnesium concentrations predominately produced transcript aborts, which reduced in proportion as the magnesium was increased as shown in Fig. 8.

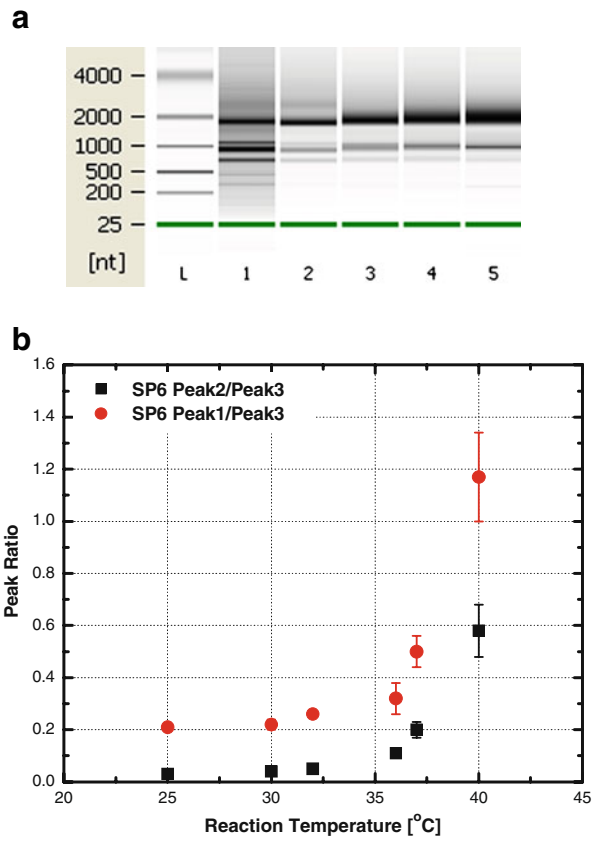

Fig. 8 Temperature effects on SP6 transcription of the AGT-SP6-H5 vRNA template using Ambion Megascript chemistry. **a** A pseudogel log plot of RNA. *Lane contents*: *L* ladder, 1–40 °C, 2–36 °C, 3–32 °C, 4–30 °C, 5–25 °C. **b** Time corrected area peak ratio of short fragments to the full-length transcript. Peak 1 is a 591-base transcript, peak 2 is a 784-base transcript, and peak 3 is the full-length 1,776-base transcript. The 37 °C data point represents average of the data from prior AGT-T7-H5 syntheses. *Error bars* represent standard deviation of multiple runs. $n=3$ at $t=60$, 120, and 240 min and $n=2$ at 360 and 960 min

Discussion

Multiple IVT Products Are the Result of Early Termination Not Self-Cleavage

Reversal of the H5 template in the transcript insert from "AGT" to "AGC" direction, shown in Fig. 1, demonstrated that the transcription sequence eliminated the production of shorter RNA sequences by early termination. With a cleavage mechanism, production of RNA bands should be independent of the transcription system. The main full-length band would appear early, followed by production of the fragment bands, which was not observed. RNA produced by various SP6 templates using the Megascript kit (Fig. 2a) showed an identical pattern of bands suggesting that template preparation was not responsible for the banding pattern. DNA templates isolated from plasmids and PCR product (Fig. 2b) produced bands of expected sizes. The PCR product, which lacks additional plasmid sequences, eliminates extraneous priming sites. RNA from Fig. 2a and reverse-transcribed in Fig. 2c shows that full-length transcripts of H5 are present. Figure 3 shows that the ratio of the bands produced over time was constant and likely eliminates cleavage as a possible source of the bands.

Following these results, the four plasmids described in Fig. 1 were fully digested and shown in Fig. 4. These sequenced templates were used in an expanded series of experiments using both NEB and Megascript chemistry. Electropherograms of Megascript products for both SP6 and T7, shown in Fig. 5, were reproducible and consistent. The

forward transcription of the vRNA form of the template (AGT clones) produced additional bands while the cRNA form (AGC clones) produced single bands. The electropherograms from Megascript transcriptions possess a small shoulder following the main peak at 1,800–2,000 that is not seen with NEB reactions. The ratio of peaks from Megascript chemistry and the AGT template are shown in Fig. 7 to be constant in time. A cleavage mechanism again was rejected. Increasing temperature in Fig. 8 yields increasing ratio of termination. Cooler temperatures minimize the transcript termination but also reduce yield (data not shown).

These templates were compared again using T7 and SP6 chemistry from NEB. Figure 9 shows the impact of increasing magnesium concentration on the termination peak ratios. NEB recommends 0.5 mM each NTP and Mg^{2+} levels be 4 mM above the total NTP concentration. A 2-mM total rNTP and 6 mM Mg^{2+} produces a high fraction of termination fragments. The vRNA AGT clone shows a minimum termination fraction above 10 mM Mg^{2+}. Termination is minimized for the cRNA AGC clone at 14 mM Mg^{2+}, a concentration

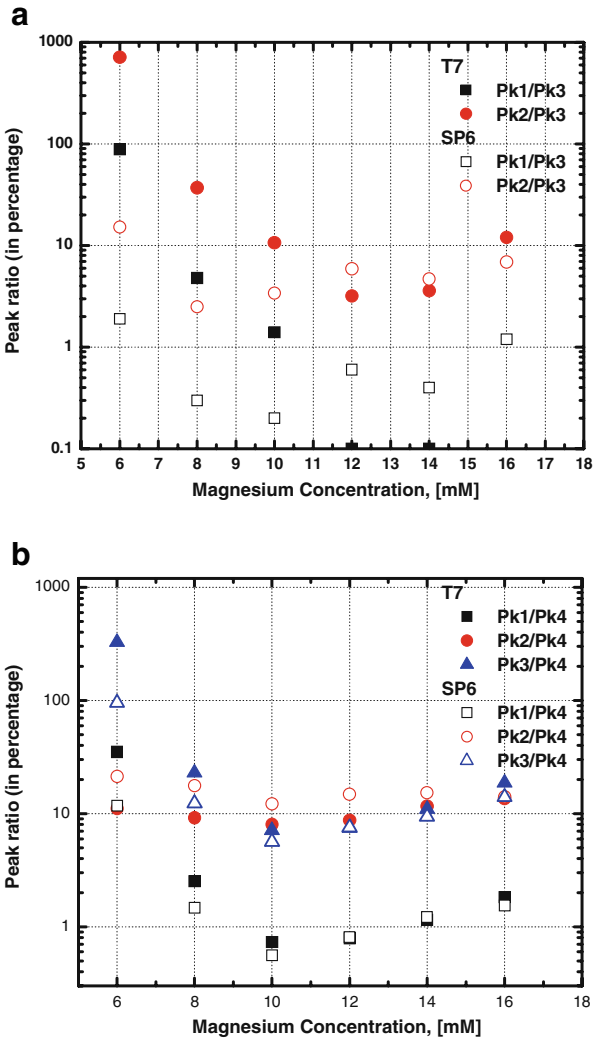

Fig. 9 Peak ratios for **a** AGT-H5 (vRNA) templates and **b** AGC-H5 (cRNA) templates

which is elevated above the recommended concentration. Since Mg^{2+} disrupts RNA secondary structure, the observed decrease in termination products with elevated Mg^{2+} validates the hypothesis that an unusually strong hairpin structure disrupts transcription.

Impact for Diagnostic RNA Synthesis

Surprisingly, the previously reported termination and pause site "ATCTGTT" did not produce the dominant termination peak in the AGT-clone reactions for SP6 and T7 reactions—instead, the major termination peak was found at base 784, as is shown in Fig. 6. The Megascript reactions for T7 and SP6, which contain proprietary concentrations of magnesium, produced enzyme-specific transcript termination profiles. Similarities of fragment profiles of both NEB and Ambion chemistries confirmed that termination is determined by the enzyme system and not the RNA polymerase kit supplier. The experiments using the NEB chemistry at adjustable Mg^{2+} concentrations suggests that the T7 Megascript formulation contain a higher concentration of Mg^{2+} than the Megascript SP6 chemistry. The directional dependence of the transcript abort also suggests that if secondary structure of the transcribed RNA triggered termination, non-symmetric downstream sequences also influence the termination event.

Stop sequences for phage transcription include type I hairpins and type II signal sequence-dependent terminator 5′ HATCTGTT 3′ in the non-template strand plus a T-rich strand. Catalytic site misalignment has been proposed as a mechanism for disrupting the normally highly processive phage polymerases at these sites [5]. The H5 influenza RNA template contains this sequence in one location in both directions. These sites are separated by more than 400 bases so a simple hairpin formation is not expected. Since RNA transcribed in the AGC orientation does not contain early termination products, the AGC-SP6-H5 template may not adopt a disruptive structure even with the termination sequence.

The issue of transcript termination applies broadly to many RNA synthesis applications. The complex nature of RNA secondary structure limits our understanding of the conditions necessary to produce full-length RNA copies. Efficient transcription of full-length products is required for many classes of assays, such as microarray analysis and NASBA techniques. NASBA assays generally limit the size of the amplified product to less than 600 bases, and many are significantly shorter [40, 41]. For the influenza genome, both ends are highly conserved as required by the native viral polymerase for a priming site, but the rest of the sequence is highly variable. Full-length transcription would lessen the need to redesign primers and only require new molecular beacons to subtype influenza samples.

Conclusions

The H5 influenza genome has particular molecular nuances that may affect lab methods and procedures. The sequencing results of the plasmid and the restriction digest, which showed clean bands, both suggest that the IVT template was not contaminated with non-target templates that would cause banding. Transcription reactions for T7 and SP6 sampled in time maintained a constant reaction ratio of full-length product to shorter transcripts, which rejects an RNA self-cleavage mechanism. The purity of RNA from H5 human influenza templates was dependent on both the direction of transcription and the RNA polymerase system. The SP6 RNA polymerase produced two short length transcripts comprising 20±3% and 50±6% of the full-length 1,776-base RNA product. The AGT-SP6 RNA solution therefore consists of 591, 784, and 1,776-base transcripts comprising 12%, 29%,

and 59% of the mixture, respectively. The T7 reaction produced a smaller fraction of early termination transcripts, which averaged only 12% of the full-length 1,776-base RNA product. A reduction in temperature reduces the early termination of these transcripts, but less RNA is synthesized.

The amplification of full-length RNA for influenza has application for viral diagnostics. Amplification using T7 and SP6 phage chemistry, while well-known, does not provide a single formulation solution. The motivation of this investigation was to determine the root cause of contaminating RNA bands in pure template mixtures and to find optimal reaction conditions. The classic type I termination sequence 5′ HATCTGTT 3′ may have caused pausing but did not significantly contribute to termination. We have shown that the extra bands are not caused by a self-cleavage mechanism representative of a ribozyme function. Magnesium proved to be the critical factor in relaxing the RNA structure such that termination products were minimized. We recommend using the T7 RNA polymerase system at double the recommended Mg^{2+} concentration when using the NEB reaction formula for the H5 influenza sequence. While this study focused exclusively on H5 influenza RNA, in vitro transcription reactions of other influenza RNA subtypes may possess similar types of hairpin sequences. Overcoming the early termination of transcription will thus be essential to generating full-length transcripts for a number of diagnostic applications.

Acknowledgments We wish to thank Dr. Tricia Serio in the Brown Biochemistry Department and Dr. Charles Vaslet in the laboratory of Dr. Agnes Kane for helpful discussions in experimental design and data interpretation. Thanks to the National Science Foundation Grant, BES-0555874, and the Brown University office of research and development for SEED funding for RNA diagnostic development.

References

1. Fung, Y. W. W., Lau, L. T., & Yu, A. C. H. (2004). The necessity of molecular diagnostics for avian flu. *Nature Biotechnology, 22*, 267.
2. Easley, C. J., Karlinsey, J. M., Bienvenue, J. M., Legendre, L. A., Roper, M. G., Feldman, S. H., et al. (2006). A fully integrated microfluidic genetic analysis system with sample-in-answer-out capability. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 19272–19277.
3. Compton, J. (1991). Nucleic-acid sequence-based amplification. *Nature, 350*, 91–92.
4. Nygaard, V., & Hovig, E. (2006). Options available for profiling small samples: A review of sample amplification technology when combined with microarray profiling. *Nucleic Acids Research, 34*, 996–1014.
5. Boudvillain, M., Schwartz, A., & Rahmouni, A. R. (2002). Limited topological alteration of the T7 RNA polymerase active center at intrinsic termination sites. *Biochemistry, 41*, 3137–3146.
6. Mukherjee, S., Brieba, L. G., & Sousa, R. (2003). Discontinuous movement and conformational change during pausing and termination by T7 RNA polymerase. *The EMBO Journal, 22*, 6483–6493.
7. Santangelo, T. J., & Roberts, J. W. (2004). Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Molecular Cell, 14*, 117–126.
8. Brooks, E. M., Sheflin, L. G., & Spaulding, S. W. (1995). Secondary Structure in the 3′-UTR of EGF and the choice of reverse transcriptases affect the detection of message diversity by RT-PCR. *Biotechniques, 19*, 806–812.
9. Chan, C. H., Lin, K. L., Chan, Y., Wang, Y. L., Chi, Y. T., Tu, H. L., et al. (2006). Amplification of the entire genome of influenza A virus H1N1 and H3N2 subtypes by reverse-transcription polymerase chain reaction. *Journal of Virological Methods, 136*, 38–43.
10. Tunitskaya, V. L., & Kochetkov, S. N. (2002). Structural–functional analysis of bacteriophage T7 RNA polymerase. *Biochemistry-Moscow, 67*, 1124–1135.
11. Ioannou, Y., Giles, I., Lambrianides, A., Richardson, C., Pearl, L. H., Latchman, D. S., et al. (2006). A novel expression system of domain I of human beta2 glycoprotein I in *Escherichia coli*. *BMC Biotechnology, 6*, 8.

12. Joshi, B. H., & Puri, R. K. (2005). Optimization of expression and purification of two biologically active chimeric fusion proteins that consist of human interleukin-13 and pseudomonas exotoxin in *Escherichia coli*. *Protein Expression and Purification, 39*, 189–198.
13. Saida, F., Uzan, M., Odaert, B., & Bontems, F. (2006). Expression of highly toxic genes in *E-coli*: Special strategies and genetic tools. *Current Protein & Peptide Science, 7*, 47–56.
14. Chen, J. M., Guo, L. X., Sun, C. Y., Sun, Y. X., Chen, J. W., Li, L., et al. (2006). A stable and differentiable RNA positive control for reverse transcription-polymerase chain reaction. *Biotechnology Letters, 28*, 1787–1792.
15. Jeong, W., & Kang, C. W. (1997). The histidine-805 in motif-C of the phage SP6 RNA polymerase is essential for its activity as revealed by random mutagenesis. *Biochemistry and Molecular Biology International, 42*, 711–716.
16. Ma, C. Q., Lyons-Weiler, M., Liang, W. J., LaFramboise, W., Gilbertson, J. R., Becich, M. J., et al. (2006). In vitro transcription amplification and labeling methods contribute to the variability of gene expression profiling with DNA microarrays. *The Journal of Molecular Diagnostics, 8*, 183–192.
17. Pokrovskaya, I. D., & Gurevich, V. V. (1994). In-vitro transcription—preparative RNA yields in analytical scale reactions. *Analytical Biochemistry, 220*, 420–423.
18. Woodrow, K. A., Airen, I. O., & Swartz, J. R. (2006). Rapid expression of functional genomic libraries. *Journal of Proteome Research, 5*, 3288–3300.
19. Li, Y., Elashoff, D., Oh, M., Sinha, U., St John, M. A. R., Zhou, X. F., et al. (2006). Serum circulating human mRNA profiling and its utility for oral cancer detection. *Journal of Clinical Oncology, 24*, 1754–1760.
20. Rodriguez-Lazaro, D., Hernandez, M., D'Agostino, M., & Cook, N. (2006). Application of nucleic acid sequence-based amplification for the detection of viable foodborne pathogens: Progress and challenges. *Journal of Rapid Methods and Automation in Microbiology, 14*, 218–236.
21. Romano, J. W., Williams, K. G., Shurtliff, R. N., Ginocchio, C., & Kaplan, M. (1997). NASBA technology: Isothermal RNA amplification in qualitative and quantitative diagnostics. *Immunological Investigations, 26*, 15–28.
22. Starkey, W. G., Millar, R. M., Jenkins, M. E., Ireland, J. H., Muir, K. F., & Richards, R. H. (2004). Detection of piscine nodaviruses by real-time nucleic acid sequence based amplification (NASBA). *Diseases of Aquatic Organisms, 59*, 93–100.
23. Morisset, D., Dobnik, D., Hamels, S., Zel, J., & Gruden, K. (2008). NAIMA: Target amplification strategy allowing quantitative on-chip detection of GMOs. *Nucleic Acids Research, 36*, e118.
24. Kerby, M. B., Freeman, S., Prachanronarong, K., Artenstein, A. W., Opal, S. M., & Tripathi, A. (2008). Direct sequence detection of structured H5 influenza viral RNA. *The Journal of Molecular Diagnostics, 10*, 225–235.
25. Toulokhonov, I., & Landick, R. (2003). The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. *Molecular Cell, 12*, 1125–1136.
26. Komissarova, N., & Kashlev, M. (1997). Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3′ end of the RNA intact and extruded. *Proceedings of the National Academy of Sciences of the United States of America, 94*, 1755–1760.
27. Hartvig, L., & Christiansen, J. (1996). Intrinsic termination of T7 RNA polymerase mediated by either RNA or DNA. *The EMBO Journal, 15*, 4767–4774.
28. Dunn, J. J., & Studier, F. W. (1983). Complete nucleotide-sequence of bacteriophage-T7 DNA and the locations of T7 genetic elements. *Journal of Molecular Biology, 166*, 477–535.
29. Sousa, R., & Mukherjee, S. (2003). T7 RNA polymerase. *Progress in Nucleic Acid Research and Molecular Biology, 73*(73), 1–41.
30. Yarnell, W. S., & Roberts, J. W. (1999). Mechanism of intrinsic transcription termination and antitermination. *Science, 284*, 611–615.
31. Sasaki, N., Izawa, M., Sugahara, Y., Tanaka, T., Watahiki, M., Ozawa, K., et al. (1998). Identification of stable RNA hairpins causing band compression in transcriptional sequencing and their elimination by use of inosine triphosphate. *Gene, 222*, 17–24.
32. Dobbins, A. T., George, M., Basham, D. A., Ford, M. E., Houtz, J. M., Pedulla, M. L., et al. (2004). Complete genomic sequence of the virulent salmonella bacteriophage SP6. *Journal of Bacteriology, 186*, 1933–1944.
33. He, B., Kukarin, A., Temiakov, D., Chin-Bow, S. T., Lyakhov, D. L., Rong, M. Q., et al. (1998). Characterization of an unusual, sequence-specific termination signal for T7 RNA polymerase. *The Journal of Biological Chemistry, 273*, 18802–18811.
34. Lyakhov, D. L., He, B., Zhang, X., Studier, F. W., Dunn, J. J., & McAllister, W. T. (1998). Pausing and termination by bacteriophage T7 RNA polymerase. *Journal of Molecular Biology, 280*, 201–213.
35. Mead, D. A., Skorupa, E. S., & Kemper, B. (1985). Single stranded DNA Sp6 promoter plasmids for engineering mutant RNAs and proteins—synthesis of a stretched preproparathyroid hormone. *Nucleic Acids Research, 13*, 1103–1118.

36. Mead, D. A., Skorupa, E. S., & Kemper, B. (1986). DNA 'blue' promoter plasmids: A versatile tandem promoter system for cloning and protein engineering. *Protein Engineering, 1*, 67–74.
37. Portugal, J., & Rodriguez-Campos, A. (1996). T7 RNA polymerase cannot transcribe through a highly knotted DNA template. *Nucleic Acids Research, 24*, 4890–4894.
38. Neumann, G., Brownlee, G. G., Fodor, E., & Kawaoka, Y. (2004). Orthomyxovirus replication, transcription, and polyadenylation. *Biology of Negative Strand RNA Viruses: The Power of Reverse Genetics, 283*, 121–143.
39. Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research, 31*, 3406–3415.
40. Collins, R. A., Ko, L. S., So, K. L., Ellis, T., Lau, L. T., & Yu, A. C. H. (2003). A NASBA method to detect high- and low-pathogenicity H5 avian influenza viruses. *Avian Diseases, 47*, 1069–1074.
41. Lanciotti, R. S., & Kerst, A. J. (2001). Nucleic acid sequence-based amplification assays for rapid detection of West Nile and St. Louis encephalitis viruses. *Journal of Clinical Microbiology, 39*, 4506–4513.