

# Kinetic Modeling and Simulation of In Vitro Transcription by Phage T7 RNA Polymerase

Sabine Arnold, Martin Siemann, Kai Scharnweber, Markus Werner, Sandra Baumann, Matthias Reuss

Institute of Biochemical Engineering, University of Stuttgart, Allmandring 31, D-70569 Stuttgart, Germany; telephone: +49-711-685-4573; fax: +49-711-685-5164; e-mail: reuss@ibvt.uni-stuttgart.de

Received 6 April 2000; accepted 15 October 2000

**Abstract:** This study provides a mathematical model of T7 RNA polymerase (T7 RNAP) kinetics under in vitro conditions targeted at application of this model to simulation of dynamic transcription performance. A functional dependence of transcript synthesis rate is derived based on: (a) essential reactant concentrations, including T7 RNAP and its promoter, substrate nucleotides, and the inhibitory byproduct inorganic pyrophosphate; (b) a distinction among vector characteristics such as recognition sequences regulating transcription initiation and termination, respectively; and (c) specific properties of the nucleotide sequence including both transcript length and nucleotide composition. Inactivation kinetics showed a half-life of T7 RNAP activity of 50 min under the conditions applied in vitro using the isolated enzyme. Model parameters and their precision are estimated using dynamic simulation and nonlinear regression analysis. The particular novelty of this model is its capability to incorporate linear genomic sequence information for simulation of nonlinear in vitro transcription kinetics. © 2001 John Wiley & Sons, Inc. *Biotechnol Bioeng* 72: 548–561, 2001.

**Keywords:** T7 RNA polymerase; transcription initiation, elongation, and termination; kinetic modeling; dynamic simulation; gene expression

## INTRODUCTION

Dynamic modeling of the transcription process is an important issue for: (i) providing an increased understanding of the overall reaction mechanism as part of gene expression regulation; (ii) application in an integrated systems approach for simulation and optimization of coupled transcription/translation; and (iii) the design of large-scale in vitro RNA production (Kern and Davis, 1999). A variety of mathematical models have been developed to elucidate transcription kinetics; however, they are typically confined to description of only some aspects of the entire mechanism.

Several studies have assumed the reaction rate to be governed primarily by initiation with no further rate-limiting

steps taking place (Martin and Coleman, 1987; Maslak and Martin, 1994; Újvári and Martin, 1996). Transcription elongation has been expressed as a ping-gong bi-bi mechanism for an analysis using the poly[d(A-T)] template and *Escherichia coli* RNA polymerase (Rhodes and Chamberlin, 1974). Rhodes and Chamberlin also indicated a competition among substrate nucleotides leading to a reduced reaction rate. The impact of conformational changes between different isomerization forms of enzyme complexes on the appearance of cooperative transcription kinetics has been investigated for wheat-germ RNA polymerase II (Job et al., 1988). T7 RNAP kinetics have been derived from a reaction scheme combining initiation, elongation, and termination (Pozhitkov et al., 1998). Using Laplace transforms, and including various simplifying assumptions, Pozhitkov's group was able to express transcription kinetics in terms of a Michaelis–Menten-type equation for each initiating and noninitiating nucleotide separately. A model for T7 RNAP kinetics has been developed for short oligonucleotides that considers the impact of both pH and free  $Mg^{2+}$  concentration on the reaction rate (Young et al., 1997). However, despite the number of models available, a satisfactory mechanistic model to depict the effects of all key reactant concentrations and, in particular, discriminating the influence of template sequence on transcript synthesis rate in one mathematical description, is yet to be fulfilled.

The relatively simple architecture of bacteriophage T7 RNA polymerase (EC 2.7.7.6) consisting of a single subunit of 98 kDa (Stahl and Zinn, 1981) and the capability of factor-independent transcription termination make the enzyme an ideal system for study of the RNA synthesis mechanism. Furthermore, T7 RNAP has been applied routinely in cell-free gene expression systems for the production of sensitive protein products, as reviewed earlier (Jerminus et al., 1998; Stiege and Erdmann, 1995). Among the issues examined, particular focus has been placed on promoter recognition as part of the initiation process (Basu and Maitra, 1986; Ikeda and Richardson, 1987; Jia and Patel, 1997; Martin and Coleman, 1987; Maslak and Martin, 1994; Sen and Dasgupta, 1993) and the occurrence of abortive transcription products (Sousa et al., 1992), which are short

Correspondence to: M. Reuss

Contract grant sponsor: German Ministry of Research

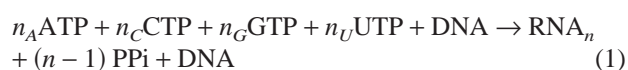
Contract grant number: A3.10U (University of Stuttgart)

Contract grant sponsor: Roche Diagnostics GmbH

Contract grant number: FKZ 0 311 302

oligonucleotides synthesized up to 8 to 14 nucleotides in length and released before entering a processive elongation mode. GTP is the initiator nucleotide that is incorporated in the growing RNA chain (Basu and Maitra, 1986; Jia and Patel, 1997; Martin and Coleman, 1987; Sen and Dasgupta, 1993). RNA synthesis can be initiated successfully by replacing GTP with its analogs, such as GMP and guanosine (Martin and Coleman, 1989). Transcript slippage might occur when the initially transcribed region starts out with GGG and when GTP is the only nucleotide present in solution (Martin et al., 1988). The termination process is understood as a reversal of the switch from an abortive to a processive transcription stage, where the interactions between the nascent RNA chain and nucleic acid binding site are disrupted, for example, through hairpin formation (Job et al., 1988; Sousa et al., 1992).

The reaction stoichiometry for net synthesis of an RNA transcript consisting of  $n$  nucleotides is given by:



where  $n_A$ ,  $n_C$ ,  $n_G$ , and  $n_U$  represent the number of adenine, cytosine, guanine, and uracil bases, respectively, in each copy of fully transcribed RNA, the sum of which equals  $n$ . The presence of inorganic pyrophosphate (PPi) was shown to inhibit the RNA synthesis reaction (Cunningham and Ofengand, 1990; Guajardo and Sousa, 1997).

Apart from the many modeling approaches available, there are only a few examples that demonstrate the full potential of biopolymer synthesis models for simulation application. The kinetic model of transcription developed by Young et al. (1997) has been applied for the optimization of the initial reaction conditions of a batch process for RNA synthesis with the aim of yield optimization. Transcription was modeled in a discrete, event-oriented approach as part of prokaryotic gene expression by Carrier and Keasling (1997). These investigators used a probabilistic model for the RNAP population to describe steric interactions, such as queueing effects among the enzymes. The applicability of discrete modeling approaches includes the ready implementation of mRNA degradation (Carrier and Keasling, 1997) and erroneous incorporation of competing substrates (Menninger, 1983). Although discrete models offer an illustrative system description, their predictive character with respect to synthesis rates is lost due to the fact that no functional relationship between reaction rates and pool concentrations is considered.

The rapid progress in modern molecular biotechniques, however, requires the development of predictive tools, particularly for use in simulation of recombinant gene expression or cell-free protein biosynthesis. Models that manage to combine transcription characteristics together with ribosomal translation and posttranslational modification steps in one model will be of great value in the optimal metabolic design of both in vitro and in vivo expression systems. In addition, the tremendous growth of biomolecular data collected worldwide offers a wealth of information for use in

bioengineering applications. Nevertheless, is great disparity between the rate and ambitiousness of data acquisition versus the actual exploitation of these libraries for mathematical modeling in process development. As such, the requirement for development of new models that allow us to make use of bioinformatic databases has been made clear (Bailey, 1998).

In the following sections, the integrative development of a kinetic model for overall in vitro transcript polymerization is illustrated, and incorporates the various facets of current understanding of the T7 RNA polymerase mechanism. A first approach considers linear genomic sequence data for use in the derived transcription model. In a systematic experimental investigation, an identification process for model constants was carried out. Using this process, in vitro polynucleotide synthesis by T7 RNAP can be simulated in terms of continuous state variable concentrations under dynamically changing conditions and for arbitrary RNA transcripts comprised of the promoters and terminators investigated.

## MATERIALS AND METHODS

### Plasmids

Three different vector systems were investigated in this study, each carrying three diverse terminator sequences. Circular plasmid pT3/T7luc was obtained from Clontech, Inc. (Palo Alto, CA). This vector contains the gene of firefly luciferase controlled by both a T3 and a T7 promoter (−17 to +6 sequence: TAATACGACTCACTATAGGGCGA), the latter of which is abbreviated  $P_{pT3/T7luc}$  herein. Among the terminators included on pT3/T7luc, only the terminator sequence of the  $\beta$ -lactamase terminator ( $T_{bla}$ ) could be identified (AAACCACCGCTGGTAGCGGTGGTTT). The other two terminator sequences, which are unknown, are subsequently referred to as  $T_{pT3/T7luc,1}$  and  $T_{pT3/T7luc,2}$ , respectively. The RNA products synthesized from pT3/T7luc were estimated from acrylamide gel analysis to have lengths of 1.89, 2.41, and 2.951 kb, respectively.

Plasmid pETK411BscFv was a kind gift from Prof. R. Schmid (Institute of Technical Biochemistry, University of Stuttgart, Germany). For the construction of pETK411BscFv, the structural gene of the recombinant single-chain Fv fragment (scFv) of antibody K411B against atrazine was taken from pCANTABscFv (Kramer and Hock, 1996) and inserted into pET20b<sup>+</sup> (Novagen, Madison, WI) (Studier and Moffat, 1986) using the restriction sites *SalI/NcoI*. Accordingly, the coding sequence of L-hydantoinase from *Arthrobacter aurescens* DSM3745 (May et al., 1998) was subcloned into pET11a (Stratagene, La Jolla, CA) using the cloning sites *NdeI/BamHI*.

Both plasmids, pEThyd and pETK411BscFv, were derived from the pET *E. coli* expression system (Studier et al., 1990) controlled by the T7 promoter,  $\phi 10$ , subsequently termed  $P_{\phi 10}$  in this study (−17 to +6 sequence given by TAATACGACTCACTATAGGGAGA). In addition to ter-

minator  $T_{\phi}$  (CAAAAACCCCTCAAGACCCGTTA-GAGGCCCAAGGGTTATGCTAG) for T7 RNAP and the  $\beta$ -lactamase terminator, another termination site was found on pETK411BscFv and pEThyd, respectively. This latter terminator is subsequently denoted  $T_{pET,1}$ , and is assumed to be identical for both plasmids, because they were derived from the same vector system. The recognition sequence of  $T_{pET,1}$  was not analyzed further. RNA transcripts of lengths 1.512, 2.406, and 3.057 kb were obtained from pEThyd, whereas pETK411BscFv led to the synthesis of 0.996-, 2.175-, and 2.744-kb RNA products.

### In Vitro Transcription Assay

The transcription buffer contained 6 mM  $MgCl_2$ , 2 mM spermidine, 10 mM DTT, and 40 mM Tris-HCl (pH 8.0). Nucleoside 5'-triphosphates (NTP) were obtained from Roche Diagnostics GmbH (Mannheim, Germany). When nonlimiting concentrations of NTP were applied, the respective final concentrations were chosen between 0.8 and 1 mM, unless otherwise indicated. The final plasmid concentration was 84 nM, unless specified otherwise. The incubation mixture was generally thermostated at a temperature of 37°C. The transcription reaction was initiated by the addition of T7 RNAP (Roche Diagnostics). The specific activity of T7 RNAP was determined in this study to be 320,000 U/mg, using maximum rate measurements as explained in what follows, together with the determination of protein amount via silver-stained sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Sambrook et al., 1989). For analysis of the Michaelis-Menten constants of GTP and UTP, yeast inorganic pyrophosphatase (PPase; Sigma-Aldrich Chemie GmbH, Deisenhofen, Germany) was added to obtain the indicated final activity. The addition of PPase to the reaction system did not lead to any RNase contamination (data not shown).

### Stability of T7 RNA Polymerase

T7 RNAP (3 U/ $\mu$ L) was incubated in a continuous-flow membrane reactor (8MC Micro-Ultrafiltration System, Amicon, Inc., Beverly, MA) filled with a reaction volume of 1 mL and operated at a constant flow rate (2 mL/h) of both feed and effluent flow. The membrane (Diaflo YM10, Amicon, Inc.) used for retention of macromolecular components such as template DNA, isolated enzyme, and product RNA had a 10-kDa cut-off. The initial composition of reactor content was adjusted according to the standard transcription protocol with the following deviations: 0.2 mM of each NTP was applied, and the reaction temperature was thermostated at 30°C. The feeding solution contained transcription buffer, as described earlier, with a modification of 0.2 mM for each NTP. At distinct timepoints, 7.5- $\mu$ L samples were taken directly from the reaction chamber and investigated immediately in initial rate studies of isolated T7 RNAP. The transcription assay (15  $\mu$ L total reaction volume) contained all of the aforementioned components; but it used

1 mM of each NTP and omitted the enzyme, which was added from the reactor experiment.

### Radioactive Labeling of RNA

Isotope labeling of the transcript was usually carried out with [5,6- $^3H$ ]-UTP exhibiting a specific activity of  $\gamma = 1.67$  TBq/mmol as noted by the manufacturer (Amersham Pharmacia Biotech, Buckinghamshire, UK). A UTP:[ $^3H$ ]-UTP ratio ranging from 360 to 720, depending on the reaction conditions chosen, was found to deliver sensitive measuring signals. The aqueous counting scintillant (ACS II) was also obtained from Amersham Pharmacia. Samples were investigated in a Tri-Carb Liquid Scintillation Analyzer (Model 1900TR, Packard Instrument Co.). Each sample of the transcription assay was precipitated with a 10% TCA solution on a Whatman filter (GC/F; 25 mm in diameter). Subsequently, the filter was rinsed with a 5% TCA solution and dried using acetone. The filter was then placed on the bottom of the scintillation vessel and submerged in 4 mL of scintillation cocktail. The samples were incubated for 1 h in complete darkness before placement into the scintillation chamber for analysis.

RNA size evaluation was performed in denaturing 5% PAGE containing 7 M urea at 65°C according to Sambrook et al. (1989). Transcription was thus carried out using [ $^{14}C$ ]-ATP (1.92 GBq/mmol, as indicated by Amersham Pharmacia) and was diluted sixfold with nonlabeled adenosine triphosphate (ATP) to achieve a final concentration of 1.2 mM. Other initial reaction conditions were: 1 mM NTP concentration; 4.5 nM of the investigated plasmid; 2 U/ $\mu$ L T7 RNAP; and the remaining conditions were as described for the standard transcription buffer. RNA molecular weight marker III (Roche Diagnostics) was taken as a reference for length determination. Prior to electrophoresis, 10- $\mu$ L samples were mixed with 10  $\mu$ L of denaturing gel loading buffer containing 7 M urea and subsequently incubated for 10 min at 65°C. Fluorograms of radiolabeled gels were performed on Hyperfilm MP (Amersham Pharmacia) following the manufacturer's protocol. Scanned fluorograms were quantified by means of the IMAGEMASTER software package (Amersham Pharmacia).

### Quantification of Radiolabeled RNA

After size determination of the observed transcripts, the concentration of each individual RNA species,  $i$ , with  $i = 1$  to  $R$  (the number of RNA species obtained for each template DNA), was quantified solely from scintillation measurements as follows:

$$C_{RNA,i} = \frac{1}{\gamma} \cdot \frac{\delta}{\alpha} \cdot \frac{m_i n_i}{n_{U,i}} \cdot X \quad (2)$$

The parameters used in Eq. (2) are conversion factors relating RNA concentration to the radioactivity signal,  $X$  (measured in dpm/ $\mu$ L). The radioactivity measured was

corrected by subtraction of the background signal determined for reactions without any plasmid. The specific activity,  $\gamma$ , of the employed radionuclide was introduced earlier.  $\delta$  represents the dilution factor between nonlabeled and [ $^3\text{H}$ ]-labeled UTP. Parameter  $\alpha$  describes the reduction of measured radioactive signal due to transcript immobilization on a Whatman filter in comparison to freely dissolved radionuclides.  $\alpha$  equaled 0.3 for [ $^3\text{H}$ ]-UTP in our measurement configuration.  $m_r$  denotes the proportion (% IOD) of RNA species  $r$  with respect to the total RNA concentration measured via acrylamide gel analysis. Constant  $n_i$  specifies the length of transcript  $i$ .  $n_{\text{U},i}$  is the number of uracil residues of transcript  $i$ . Throughout this study, RNA concentration is expressed in nucleotide moles per reaction volume, in all cases. The sum of the individual RNA species synthesized from this plasmid comprises the total RNA concentration,  $C_{\text{RNA}}$ .

## Computational Methods

Kinetic modeling was assisted by an automated algorithm (Mauch et al., 1997) written in MAPLE V (release 4, Waterloo Maple, Inc.), which was developed specifically for kinetic model-building based on the pseudo-steady-state assumption (Cleland, 1963) and partial rapid equilibrium treatment (Cha, 1968).

Simulation studies and parameter identification were performed on a SUN-workstation using the numerical integrator, ACSL (version 11.5.3, Advanced Continuous Simulation Language, Mitchell and Gauthier Associates, Inc. [Gear algorithm]), in combination with the optimization package, OPTDESX (version 2.0.4, Design Synthesis, Inc. [Simulated annealing algorithm]). When parameter precision is given in this study, 95% confidence interval computations were executed with software (Merkel et al., 1996) developed on the basis of a previously described linear approximation method (Draper and Smith, 1981) or by applying Student's  $t$ -distribution.

The length of transcribed regions for T7 RNAP were determined using the pattern recognition program, TERMINATOR (Brendel and Trifonov, 1984), run with the WISCONSIN package (Genetic Computer Group) for identification of termination sequence.

Free  $\text{Mg}^{2+}$  ion concentrations were calculated by an algorithm written according to a previously proposed method (Storer and Cornish-Bowden, 1976), which is based on iteratively solving ionic species mass balance equations. The dissociation constants applied in this method for complex formation between dissolved ions and low-molecular-weight system components were taken from Langer et al. (1977) unless otherwise noted:  $K_{\text{HNTP}} = 1.26 \cdot 10^{-7} \text{ M}$ ,  $K_{\text{H}_2\text{NTP}} = 8.71 \cdot 10^{-5} \text{ M}$ ,  $K_{\text{Mg}_2\text{NTP}} = 5.01 \cdot 10^{-2} \text{ M}$ ,  $K_{\text{MgNTP}_2} = 3.98 \cdot 10^{-3} \text{ M}$ ,  $K_{\text{MgNTP}} = 1.15 \cdot 10^{-5} \text{ M}$  (O'Sullivan and Perrin, 1964),  $K_{\text{MgHNTP}} = 1.41 \cdot 10^{-3} \text{ M}$  (O'Sullivan and Perrin, 1964),  $K_{\text{MgPi}} = 1.32 \cdot 10^{-2} \text{ M}$  (Smith and Alberty, 1956),  $K_{\text{MgPPi}} = 1.70 \cdot 10^{-5} \text{ M}$  (Käpylä et al., 1995), and  $K_{\text{Mg}_2\text{PPi}} = 2.14 \cdot 10^{-3} \text{ M}$  (Käpylä et al.,

1995). The employed buffer, Tris, was shown to exhibit no significant binding capacity to metal ions (Good et al., 1966). Magnesium bound with nucleic acids was accounted for through a rough estimate, assuming an average association of 1 molecule of  $\text{Mg}^{2+}$  per each two phosphate groups of nucleic acid (Record et al., 1976). When free  $\text{Mg}^{2+}$  ion concentrations are given in this study, they represent the geometric mean between the respective concentrations calculated at the start and termination of each experiment.

## RESULTS

### Reaction Kinetics

The rate equation for overall transcript polymerization by T7 RNA polymerase has been derived from the reaction scheme (Fig. A1) given in the Appendix. Rate derivation was assisted by an algorithm that had been developed for the symbolic computation of general enzymatic rate equations using MAPLE V (Mauch et al., 1997). MAPLE V is particularly useful in dealing with complex reaction mechanisms consisting of a large number of individual reaction steps, such as those encountered in overall polymerization kinetics. With the introduction of lumped parameters, and after suitable rearrangement, the rate of in vitro transcription by T7 RNAP is:

$$v = \frac{V_{\text{max}}}{1 + \sum_{j=1}^N \frac{K_{\text{M,NTP}_j}}{C_{\text{NTP}_j}} \left( 1 + \frac{C_{\text{PPi}}}{K_{\text{L,PPi}}} + \sum_{i \neq j}^N \frac{C_{\text{NTP}_i}}{K_{\text{L,NTP}_i}} \right) + \frac{K_{\text{M,D}}}{C_{\text{D}}} \left[ 1 + \frac{K_{\text{G}}^I}{C_{\text{GTP}}} \left( 1 + \frac{C_{\text{PPi}}}{K_{\text{L,PPi}}} + \sum_{i=1}^{N-1} \frac{C_{\text{NTP}_i}}{K_{\text{L,NTP}_i}} \right) \right]} \quad (3)$$

Eq. (3) shows the impact of concentrations of nucleoside triphosphates ( $C_{\text{NTP}}$ ), total promoter ( $C_{\text{D}}$ ), and inhibition of inorganic pyrophosphate ( $C_{\text{PPi}}$ ) on total transcript synthesis rate,  $V$ .  $N$  represents the number of different ribonucleotide species (usually 4), of which the transcript product is composed. The term in square brackets in the denominator of Eq. (3) includes all the information about the initiation process: T7 RNAP binding to the promoter is modulated by initial GTP binding, inorganic pyrophosphate inhibition, and competition among substrate nucleotides ( $\text{NTP} \neq \text{GTP}$ ).

Contained in the kinetic parameters of the derived rate equation is the information about nucleotide sequence characteristics of the DNA template employed, which covers: (a) transcript length; (b) nucleotide composition; and (c) the rate constants for transcription initiation, elongation, and termination. The mathematical relation of these lumped parameters is given in Table I.

The dissociation constant for initial GTP binding,  $K_{\text{G}}^I$ , has been reported to be 25 nM, by Sen and Dasgupta (1993). The inhibition constant for ATP competing against GTP incorporation using the poly(dC) template was measured for

**Table I.** Estimated kinetic parameters for in vitro transcription by T7 RNA polymerase using plasmid pT3/T7luc and a comparison to literature data.

Parameter	Unit	This study	Chamberlin and Ring (1973)	Ikeda and Richardson (1987)
$V_{\max} = k_{\text{eff}} C_{E,t}$	( $\mu M/\text{min}$ )	$188 \pm 100$		
$K_{M,D} = \frac{k_{\text{eff}}}{k_I} K_D$	(nM)	$6.3 \pm 2.6$		
$K_{M,ATP} = n_A \frac{k_{\text{eff}}}{k_E} K_A$	( $\mu M$ )	$76 \pm 22$	47	31
$K_{M,CTP} = n_C \frac{k_{\text{eff}}}{k_E} K_C$	( $\mu M$ )	$34 \pm 7$	81	23
$K_{M,GTP} = (n_G - 1) \frac{k_{\text{eff}}}{k_E} K_G + \frac{k_{\text{eff}}}{k_I} K_G^I$	( $\mu M$ )	$76 \pm 12$	160	190
$K_{M,UTP} = n_U \frac{k_{\text{eff}}}{k_E} K_U$	( $\mu M$ )	$33 \pm 6$	60	40
$K_{I,PPi}$	( $\mu M$ )	$200 \pm 45$		
$k_d$	( $\text{min}^{-1}$ )	0.014		

T7 RNAP to be 4.8 mM (Guajardo and Sousa, 1997). Therefore, the effect of competing nucleotide substrates may be significant only at concentrations above several millimolar of competing NTP and is neglected for the remainder of this study. Omission of nucleotide competition reduces the kinetic rate equation to:

$$V = \frac{V_{\max}}{1 + \sum_{j=1}^N \frac{K_{M,NTP,j}}{C_{NTP,j}} \left( 1 + \frac{C_{PPi}}{K_{I,PPi}} \right) + \frac{K_{M,D}}{C_D} \left[ 1 + \frac{K_G^I}{C_{GTP}} \left( 1 + \frac{C_{PPi}}{K_{I,PPi}} \right) \right]} \quad (4)$$

At excess concentrations of all other nucleotides except for the  $j$ th NTP, and at saturation of promoter concentration, the rate equation is determined by:

$$V = \frac{V_{\max} C_{NTP,j}}{K_{M,NTP,j} \left( 1 + \frac{C_{PPi}}{K_{I,PPi}} \right) + C_{NTP,j}} \quad (5)$$

In the presence of sufficient inorganic pyrophosphatase activity in the reaction system, the term describing competitive inhibition vanishes in Eq. (5), leading to:

$$V = \frac{V_{\max} C_{NTP,j}}{K_{M,NTP,j} + C_{NTP,j}} \quad (6)$$

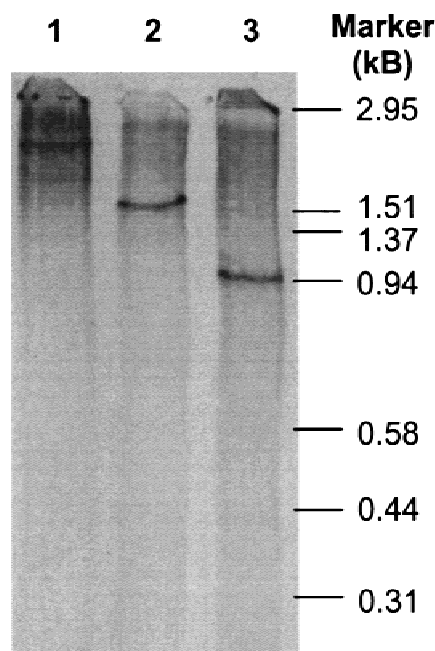
Saturation of all nucleotides simplifies Eq. (4) to a dependence of transcription rate solely on promoter concentration according to irreversible Michaelis–Menten kinetics:

$$V = \frac{V_{\max} C_D}{K_{M,D} + C_D} \quad (7)$$

### RNA Distribution

To compute RNA concentration from scintillation data, a systematic investigation of RNA length distribution was undertaken for the three plasmids, pT3/T7luc, pEThyd, and pETK411BscFv. The results of this analysis are shown in

Figure 1. Using the information of both RNA marker and theoretical RNA lengths to be obtained from nucleotide sequence analysis, all bands of RNA displayed on the gel could be quantified and assigned to the respective transcription terminators. The data evaluated are presented in Table II. From a practical standpoint, a categorization of RNA into classes of distinct lengths was desirable, although some smear of RNA bands was noted, particularly with vector



**Figure 1.** RNA length distribution investigated with acrylamide gel electrophoresis for templates (1) pT3/T7luc, (2) pEThyd, and (3) pETK411BscFv. Independent transcription assays were made under standard reaction conditions adding the respective plasmid. Ten-microliter samples were taken after 90 min of transcription reaction and treated as described. The gel was subsequently quantified using the IMAGEMASTER software (Amersham Pharmacia). The data evaluated are summarized in Table II.

**Table II.** Distinction of effective rate constants according to the various RNA species obtained for the vector systems used as well as transcript length, respective number of uracil bases, and relative amounts of individual RNA synthesized.

Plasmid	Promoter	Terminator	RNA (kb)	$n_U$	$m$ (% IOD)	$k_{\text{eff}}$ ( $\text{s}^{-1}$ )
pT3/T7luc	$P_{\text{pT3/T7luc}}$	$T_{\text{pT3/T7luc,1}}$	1.890	504	9	8.8
	$P_{\text{pT3/T7luc}}$	$T_{\text{pT3/T7luc,2}}$	2.410	634	53	52.1
	$P_{\text{pT3/T7luc}}$	$T_{\text{bla}}$	2.951	749	38	37.5
pEThyd	$P_{\phi 10}$	$T_{\phi}$	1.512	292	64	73.8
	$P_{\phi 10}$	$T_{\text{pET,1}}$	2.406	570	6	6.9
	$P_{\phi 10}$	$T_{\text{bla}}$	3.057	708	30	34.6
pETK411BscFv	$P_{\phi 10}$	$T_{\phi}$	0.996	231	68	67.6
	$P_{\phi 10}$	$T_{\text{pET,1}}$	2.175	593	4	4.0
	$P_{\phi 10}$	$T_{\text{bla}}$	2.744	718	28	27.8

pT3/T7luc. With this strategy, transcript synthesis delivered three RNA species of different lengths for each of the plasmids tested. No significant amount of read-through products was observed for either plasmid. The respective portion of any RNA species as part of total RNA synthesized in each experiment is denoted by  $m$  and given in percentage of integrated optical density (IOD). With known length of each transcript, the number of uracil residues contained in the RNA was determined from each nucleotide sequence. The values given in Table II, together with Eq. (2), formed the basis for all conversions of scintillation measurements into concentrations of both individual RNA species and total RNA concentration.

### Material Balance Equations

The following material balances apply for each of the four NTP substrates, inorganic pyrophosphate, and total RNA consisting of  $R$  different RNA species in an in vitro transcription assay using isolated T7 RNAP under batch conditions:

$$\frac{dC_{\text{RNA}}}{dt} = \sum_{i=1}^R V_i \quad (8)$$

$$\frac{dC_{\text{NTP},j}}{dt} = - \sum_{i=1}^R f_{j,i} n_i V_i \quad \text{for } j = 1 \text{ to } N \quad (9)$$

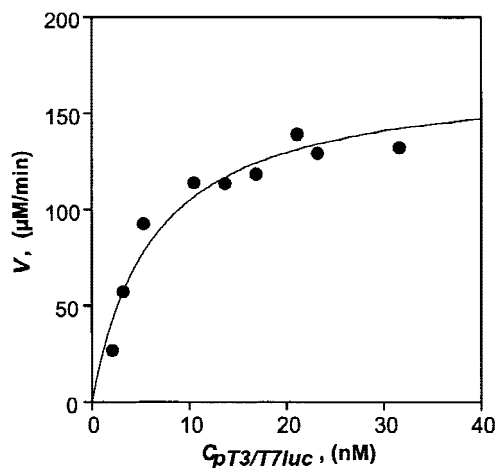
$$\frac{dC_{\text{PPi}}}{dt} = \sum_{i=1}^R \frac{n_i - 1}{n_i} V_i \quad (10)$$

where  $f_{j,i}$  denotes the percentage portion of base  $j$  contained in the  $i$ th product RNA. For each of the reaction rates used in the ordinary differential Eqs. (8)–(10), one of the rates given in Eqs. (4)–(6) was substituted, according to the system conditions investigated. Activities of DNases, RNases, and NTPases were not detectable in the in vitro transcription system examined (data not shown), but would enter the balance equations if necessary. Eqs. (8)–(10) were numerically integrated to deliver simulated time-dependent courses of substrate and product concentrations of T7 RNAP catalysis.

### Maximum Transcription Rate

The maximum transcription rate for total RNA synthesis was estimated for plasmid pT3/T7luc, applying initial rate studies. In this reaction system, all substrate nucleotides were kept at excess levels and a final T7 RNAP activity of  $1 \text{ U}/\mu\text{L}$  was applied. The rates thus received versus the varying plasmid concentrations are depicted in Figure 2. Using nonlinear regression, Eq. (7) was fitted to the experiment. The average maximum transcription rate estimated within three consecutive experiments (one of three data sets shown in Fig. 2) was  $V_{\text{max}} = 188 \pm 100 \mu\text{M}/\text{min}$ . With a known T7 RNAP concentration, the effective rate constant for overall transcription was calculated as  $k_{\text{eff}} = 97 \pm 52 \text{ s}^{-1}$  using the relation given in Table I. In other words, overall T7 RNAP progression occurred in the system examined at an average rate of 97 nucleotides per second.

Substrate affinity for T7 RNAP to promoter  $P_{\text{pT3/T7luc}}$  was determined to be  $K_{\text{M,D}} = 6.3 \pm 2.6 \text{ nM}$  from the same experimental investigation. Because this value compares favorably to previous estimates of the Michaelis–Menten constant for binding to the  $\phi 10$  promoter ranging between 6 and 8 nM (Ikeda et al., 1992), the affinity constant for promoter association found in this study was also applied for  $P_{\phi 10}$  carried on pEThyd and pETK411BscFv in an experimental



**Figure 2.** Dependence of transcription rate on concentration of plasmid pT3/T7luc in the presence of 3.7 mM free  $\text{Mg}^{2+}$  ions.

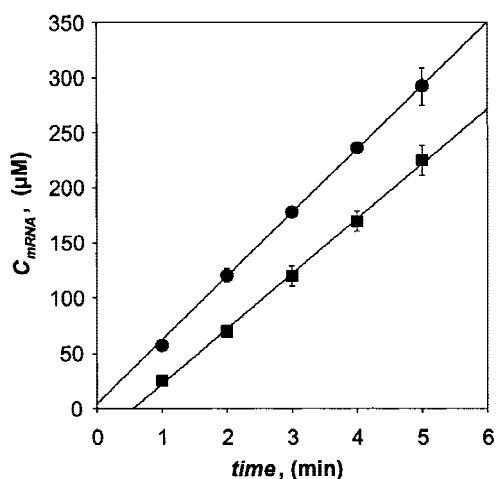
setup where the template concentration was below saturation (Fig. 3). From the data displayed in Figure 3, maximum transcription rates for T7 RNAP were estimated for plasmids pEThyd and pETK411BscFv to be  $66.5 \pm 5.0 \mu\text{M}/\text{min}$  and  $57.5 \pm 3.8 \mu\text{M}/\text{min}$ , respectively, by applying  $0.3 \text{ U}/\mu\text{L}$  of the enzyme to both reaction assays. The overall rate constant for transcription using pEThyd was  $k_{\text{eff}} = 115 \pm 9 \text{ s}^{-1}$ , whereas  $k_{\text{eff}}$  was found to be  $99 \pm 7 \text{ s}^{-1}$  for pETK411BscFv.

The overall rate constants received compare to an average T7 RNAP progression of around  $60 \text{ s}^{-1}$ , as reported for transcription of T7 DNA (Ikeda and Richardson, 1987). A rough estimate for the rate constant of chain growth using T7 RNAP was found to be 100 nucleotides per second (Chamberlin and Ring, 1973). An elongation rate constant of 230 nucleotides per second was given earlier using the T7 DNA template (Golomb and Chamberlin, 1974). As can be seen, there is a wide range of rate constants for T7 RNAP for the various template systems employed.

Similarly, the effective rate constants obtained in this study for the three different plasmids show a noticeable variation, albeit small, from 97 to  $115 \text{ s}^{-1}$ . However, the highly precise estimates of  $k_{\text{eff}}$  for the pET-derived vectors with merely overlapping 95% confidence intervals suggest a statistical significance for the observed discrepancy. Deviations may be explained by the heterogeneity of RNA products of different lengths (see Table II), as well as with differences in binding strengths and interactions with the respective initiation and termination sites.

### Rate Constants for Initiation, Elongation, and Termination

When all substrate concentrations are nonlimiting, the rate of total RNA synthesis comprises the sum of the  $R$  maxi-



**Figure 3.** Time course of total RNA concentration using plasmids pEThyd (●) and pETK411BscFv (■) at a concentration of 41 nM. Free  $\text{Mg}^{2+}$  ion concentration was 3.7 mM in each reaction assay. The respective straight line represents the average result of linear regression for the corresponding three data sets. Error bars denote the standard deviation among the respective measurements. Maximum transcription rates were calculated using Eq. (7).

imum rates for the formation of individual RNA species according to:

$$V = V_{\text{max}} = \sum_{i=1}^R V_{\text{max},i} = \sum_{i=1}^R k_{\text{eff},i} C_{\text{E},t} = k_{\text{eff}} C_{\text{E},t} \quad (11)$$

In this case, the overall rate constant for total RNA synthesis,  $k_{\text{eff}}$ , can be decomposed into the single contributions,  $k_{\text{eff},i}$ , calculated for each RNA species,  $i$ . In the course of rate derivation,  $k_{\text{eff},i}$  was found to obey:

$$k_{\text{eff},i} = \frac{1}{\frac{1}{k_i} + \frac{n_i - 1}{k_E} + \frac{1}{k_{T,i}}} \quad (12)$$

where index  $i$  denotes the parameters specifically related to RNA product  $i$ . The significance of Eq. (12) arises from the fact that it is thus possible to describe the rate of overall transcript polymerization in terms of transcript length and, importantly, by the characteristic, template-specific initiation and termination properties, respectively. The left-hand side of Eq. (12) can be easily obtained as:

$$k_{\text{eff},i} = m_i k_{\text{eff}} \quad (13)$$

and with  $m_i$  known from PAGE analysis. Computed results are given in Table II. For example, the overall rate constants for T7 RNAP progression using vector pEThyd were thus determined to be  $73.8 \text{ s}^{-1}$  for RNA of length 1.512 kb,  $6.9 \text{ s}^{-1}$  for the 2.406-kb transcript, and  $34.6 \text{ s}^{-1}$  for 3.067-kb RNA. The sum of these rate constants equals the overall rate constant for total RNA synthesis of  $115 \text{ s}^{-1}$  given earlier for this template. In other words, this means that the synthesis of the 1.512-kb RNA contributes an average of 73.8 nucleotides per second to the observed rate of total transcript synthesis using pEThyd, whereas 6.9 nucleotides per second are due to the formation of the 2.406-kb RNA, and so on.

Substitution of the respective rate constant,  $k_{\text{eff},i}$ , for all RNA species from Table II into Eq. (12) leads to a total set of nine nonlinear algebraic equations with the eight unknowns: the rate constants for initiation,  $k_{I,\text{pT3/T7luc}}$  and  $k_{I,\phi10}$ ; the elongation rate constant,  $k_E$ ; and the rate constants for the transcription terminators examined,  $k_{T,\phi}$ ,  $k_{T,\text{bla}}$ ,  $k_{T,\text{pT3/T7luc},1}$ ,  $k_{T,\text{pT3/T7luc},2}$ , and  $k_{T,\text{pET},1}$ . A successive procedure for estimation of these rate constants was carried out by minimizing the sum of squared relative errors. At first, the parameters associated with the pET-derived vectors were estimated separately. The initiation rate constant for promoter  $\phi10$  was determined to be  $k_{I,\phi10} = 79 \text{ s}^{-1}$ , together with the rate constant for elongation,  $k_E = 5.8 \cdot 10^7 \text{ s}^{-1}$ . The rate constants for termination were  $k_{T,\phi} = 630 \text{ s}^{-1}$  for the T7 terminator  $T_\phi$ , and  $k_{T,\text{bla}} = 50 \text{ s}^{-1}$  for T7 RNAP terminating at the  $\beta$ -lactamase terminator.  $k_{T,\text{pET},1}$  equaled  $5 \text{ s}^{-1}$  for the unidentified terminator sequence located on both vectors pEThyd and pETK411BscFv. Subsequently, keeping these parameters fixed, the remaining initiation and termination constants for plasmid pT3/T7luc were calculated

to be  $k_{i,pT3/T7luc} = 146 \text{ s}^{-1}$ ,  $k_{T,pT3/T7luc,1} = 80 \text{ s}^{-1}$ , and  $k_{T,pT3/T7luc,2} = 9 \text{ s}^{-1}$ , respectively.

Apart from the observed variation among the estimated parameters, the effective rate constant for T7 RNAP transcription was shown to be governed primarily by the specific initiation and termination features of the vector system applied. In general, it appears plausible that the rate constant for elongation is found to be much higher, even by several orders of magnitude, than the rate constants for initiation and termination, where there are, in both latter cases, multiple interactions required between the enzyme and the respective recognition sites. The impact of elongation rate may be significant only at extreme transcript lengths.

A comparison of the different termination rate constants shows a range of estimated values within about two orders of magnitude, from  $5 \text{ s}^{-1}$  for  $T_{pET,1}$  to  $630 \text{ s}^{-1}$  for  $T_{\phi}$ . This finding may be surprising at first, but can be explained easily. Parameter  $k_T$ , which governs the rate of release of the corresponding RNA product, is affected both by the temporary hold-up of RNA polymerase during the step of RNA dissociation and, importantly, by the competing reaction of forward progression of RNA polymerase (von Hippel and Yager, 1991). The greater the degree to which the transcribing enzyme fails to recognize the termination signal as such and simply continues to elongate, the lower the termination rate constant of our model will be. Thus, small rate constants for transcription termination cannot necessarily be attributed to a slow progression of elongating RNA polymerase. They may also be indicative of different levels of termination efficiency among the RNA transcripts compared.

Termination efficiency is defined as the fraction of RNA corresponding to this terminator with respect to the sum of all termination products synthesized from the same template DNA (Lyakhov et al., 1998; von Hippel and Yager, 1991). The efficiency of transcription termination thus specifies the relative frequency of termination events that lead to the synthesis of the RNA specific to this termination site. The comparatively high rate constant of  $630 \text{ s}^{-1}$  for termination at terminator  $T_{\phi}$ , with regard to the other termination sequences investigated, is in agreement with a high termination efficiency, which has been identified to be 80% for this terminator (Lyakhov et al., 1998). This reported termination efficiency is in accordance with our own observations, which are detailed in the Appendix. While the T7 promoter of vector pT3/T7luc appears to be more efficient than  $P_{\phi10}$ , this result should be considered with caution due to the increased measurement uncertainty associated with this parameter.

Using the rate constants identified for initiation, elongation, and termination, the expected maximum rate of transcript formation at a given concentration of T7 RNA polymerase may also be predicted for the initiation and termination sequences investigated when they are contained on other vector systems.

## Time-Dependent Catalytic Activity

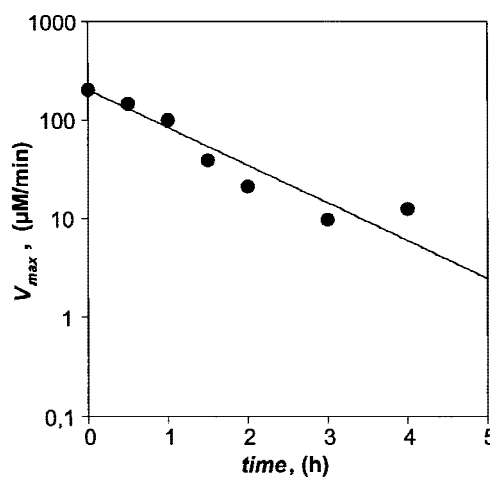
Investigation of dynamic reaction performance requires knowledge of the inactivation kinetics of T7 RNAP over the process time. As previously noted (Chamberlin and Ring, 1973; Maslak and Martin, 1994), T7 RNAP exhibits poor catalytic stability in dilute systems, which may be overcome through the addition of bovine serum albumin (BSA) to the reaction system. However, to exclude an assay contamination due to RNase, DNase, or protease activity, BSA was deliberately omitted in our experimental set-up. To determine the stability behavior of actively transcribing T7 RNAP, the enzyme was incubated under static conditions in a continuous-flow reactor containing NTP, pT3/T7luc, and transcription buffer, as described earlier. At various time-points, samples were taken from the incubation mixture, and activity measurements were started by the addition of transcription buffer containing no further enzyme. The obtained variation of initial reaction rate of isolated T7 RNAP is shown in Figure 4. The stability behavior of the enzyme can be expressed by a time-dependence of maximum reaction rate, for which a first-order reaction was fitted with linear regression:

$$V_{\max}(t) = V_{\max}^0 e^{(-k_d t)} \quad (14)$$

The resulting degradation constant was estimated to be  $k_d = 0.014 \text{ min}^{-1}$ , which translates into a half-life of T7 RNAP activity of 50 min under the reaction conditions employed. Eq. (14) was substituted into the respective rate Eqs. (4)–(6) for dynamic simulation.

## Substrate Affinity and Product Inhibition

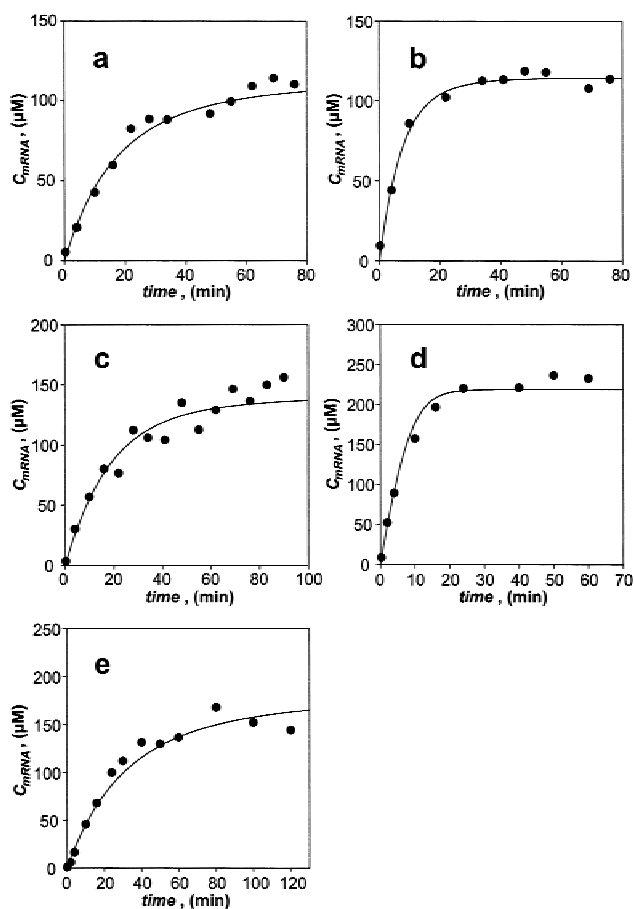
Apparent Michaelis–Menten constants were estimated from dynamic experiments where the nucleotide under investiga-



**Figure 4.** Inactivation kinetics of T7 RNAP incubated under static reaction conditions. The buffered transcription assay contained 0.2 mM NTP and an initial T7 RNAP concentration of 3 U per microliter of reaction volume. Free  $\text{Mg}^{2+}$  ion concentration was 4.9 mM in the continuous-flow reactor and 1.7 mM during the initial rate measurements. Reduction of maximum activity over the time was fitted to a first-order reaction using linear regression.

tion was added initially at a limiting concentration. In these experiments, which are displayed in Figure 5, all other substrates were kept at saturating levels over the course of the reaction. The affinity constants for ATP, CTP, GTP, and UTP were estimated to have the comparatively similar values of  $76 \pm 22$ ,  $34 \pm 7$ ,  $76 \pm 12$ , and  $33 \pm 6 \mu\text{M}$ , respectively. Although an increased  $K_M$  value was reported earlier for nucleotide GTP (cf. Table I), this finding could not be confirmed with our results.

The inhibition constant due to inorganic pyrophosphate



**Figure 5.** Comparison of the experimentally obtained and simulated time-course of total RNA concentration under batchwise conditions applying a single limiting substrate nucleotide. The concentration of the limiting NTP was  $30 \mu\text{M}$ , except for in experiments (d) and (e), where it was  $50 \mu\text{M}$ . All other substrates were held at saturating levels throughout the experiments. The resulting free  $\text{Mg}^{2+}$  concentration was 2.8, 3.2, 2.7, 2.7, and  $0.2 \text{ mM}$  for reaction conditions (a)–(e), respectively. (a) ATP limitation. The simulated curve uses parameters estimated from three consecutive experiments. (b) CTP limitation. The concentration of T7 RNAP was  $0.2 \text{ U}/\mu\text{L}$ . (c) GTP limitation. Initial concentrations applied were  $60 \text{ nM}$  of plasmid pT3/T7luc,  $0.2 \text{ U}/\mu\text{L}$  T7 RNAP, and  $0.01 \text{ U}/\mu\text{L}$  inorganic pyrophosphatase. Parameters used for simulation were estimated simultaneously from two independent experiments carried out under the same reaction conditions. (d) UTP limitation in a system containing  $0.3 \text{ U}/\mu\text{L}$  T7 RNAP and  $0.01 \text{ U}/\mu\text{L}$  inorganic pyrophosphatase. Parameters were identified simultaneously on the basis of two experiments carried out independently. (e) UTP limitation in a system with an initial addition of  $3 \text{ mM}$  PPI and  $0.6 \text{ U}/\mu\text{L}$  T7 RNAP. The inhibition constant for PPI was estimated from the data presented and held fixed in the estimation of affinity constants for substrate nucleotides when no PPase was added.

**Table III.** Estimated rate constants for transcription initiation, elongation, and termination using plasmids pT3/T7luc, pEThyd, and pETK411BscFv.

Parameter	Unit	Value
$k_{i,pT3/T7luc}$	( $\text{s}^{-1}$ )	146
$k_{i,\phi 10}$	( $\text{s}^{-1}$ )	79
$k_E$	( $\text{s}^{-1}$ )	$5.8 \cdot 10^7$
$k_{T,\phi}$	( $\text{s}^{-1}$ )	630
$k_{T,bla}$	( $\text{s}^{-1}$ )	50
$k_{T,pT3/T7luc,1}$	( $\text{s}^{-1}$ )	80
$k_{T,pT3/T7luc,2}$	( $\text{s}^{-1}$ )	9
$k_{T,pET,1}$	( $\text{s}^{-1}$ )	5

concentration was estimated to be  $200 \pm 45 \mu\text{M}$ , which is within one order of magnitude of the  $830 \mu\text{M}$  value given previously for this parameter (Guajardo and Sousa, 1997). Thus, the adverse impact of inorganic pyrophosphate is suggested to be relevant under in vitro conditions using isolated T7 RNA polymerase, where byproduct PPI is accumulated simultaneously to transcription. However, under typical conditions encountered in vivo, where, for instance, exponentially growing cells of host *E. coli* contained PPI concentrations of about  $0.5 \text{ mM}$  (Kukko-Kalske et al., 1989), PPI inhibition should play only a minor role in determining transcription rate. The same holds for systems using T7 RNAP in cellular extracts, as long as a sufficient PPase activity is maintained.

## DISCUSSION

This study has succeeded, for the first time, in developing a rigorous mechanistic representation of RNA synthesis by T7 RNA polymerase on the basis of genomic sequence information. The transcription kinetics derived in vitro aimed to incorporate the current knowledge of the RNA synthesis mechanism available, at the same time trying to balance the need for a detailed system description with a reasonable model complexity. Among the lumped model parameters used, a closer examination was carried out for the sequence-oriented parameter,  $V_{\max}$ , the maximum rate of transcript formation. In this approach, particular focus was placed on defining a functional relationship of RNA synthesis rate to transcript length and discriminating between different initiation and termination sequences. For an accurate estimation, however, of the other kinetic model constants used in the derivation of the transcription rate (e.g., the association constants of substrate nucleotides), further experiments are necessary, which are beyond the scope of this work. The model developed can be used for dynamic simulation of transcript formation using the initiation and termination sites investigated.

The derived kinetic expression is, strictly speaking, only applicable when a pseudo-steady-state synthesis is reached. This is the case when the number of initiating polymerases approximately equals the number of polymerases being released from the template. Pre-steady-state kinetics are ne-

glected in this model. However, the resulting error may be small when no polymerase queuing takes place and is further reduced with shorter transcript length. This is in agreement with a previous observation, where run-off, fall-off, and abortive transcript concentrations were found to increase linearly with the synthesis time and in approximately the same proportion (Guajardo et al., 1998). These findings suggest a steady-state synthesis by T7 RNAP to be reached within a few seconds at nonlimiting substrate conditions, supporting the applicability of the pseudo-steady-state assumption made inherently in our model derivation.

Individual rate constants for initiation, elongation, and termination were determined, making a distinction between the vector systems applied. Although further experimental evidence is needed to cover a broader set of different types of plasmid constructs, this study demonstrated a quantitatively significant impact of different initiation and termination sites on the rate of total RNA synthesis. In general, the information with regard to the types of promoters and terminators used in the experimental analyses has not been provided. Evidently, however, the efficiency of initiation and termination steps during the transcription process will vary greatly with the strength of bond formation to the respective recognition sequence. A sufficiently detailed approach providing such information will contribute significantly to clarification and allow for a more reliable comparison of experimental results among the different studies.

The impact of abortive transcription was not particularly accounted for in this study. The synthesis of abortive versus full-length transcripts was found to occur at approximately the same order of magnitude when molarities were considered (Guajardo et al., 1998; Young et al., 1997). Thus, the larger the size of the full-length transcript, the less important the relative amount of nucleotides incorporated into abortive transcription products. For the long transcripts treated in our study (on the order of 1000 bases), abortive transcription and its influence on nucleotide consumption may be insignificant, and its neglect appears reasonable. For simulating the synthesis of short transcripts, the effect of abortive RNA product formation can contribute significantly to the material balance of substrate nucleotides and should be taken into account (Young et al., 1997).

Transcript slippage, as well as the effects of further conformational steps within the reaction mechanism, were neglected in the given model structure. However, when these phenomena are a consequence of characteristic elements of the nucleotide sequence, they may be regarded as inherent in the estimates of the overall rate constants for transcription initiation, elongation, and termination, respectively.

Although apparent affinity constants for substrate nucleotides have been estimated for T7 RNAP in previous studies (Chamberlin and Ring, 1973; Ikeda and Richardson, 1987), the therein applied conditions suggest some caution in the interpretation of reported results. One restriction associated with these studies arises from the use of initial substrate concentrations that may have been insufficient to guarantee substrate saturation. Common to these earlier works is the

analysis of initial rate measurements with graphical methods such as Lineweaver–Burk or Eadie–Hofstee plots. Unfortunately, however, the graphical analysis itself was not provided in either of the aforementioned studies, so an assessment of measurement uncertainty contained within the estimation results was eluded. Furthermore, the addition of BSA to these systems might have introduced additional reactions involving nucleotide substrates or those potentially causing template and product degradation. An investigation of genuine T7 RNAP kinetics, however, requires examination of the isolated enzyme *in vitro* without superposition of other reactions.

Apparent Michaelis–Menten constants for nucleotides ATP, CTP, and UTP were estimated in this study within an order of magnitude of previous data, ranging from 23 to 81  $\mu\text{M}$  (compare with the referenced data in Table I). One notable exception to the literature data is the affinity constant for GTP, which was estimated to be 76  $\mu\text{M}$  from our experiments. In contrast, a raised  $K_M$  value, ranging from 160 to 190  $\mu\text{M}$ , has been typically found for GTP when using bacteriophage T7 RNA polymerase (Chamberlin and Ring, 1973; Ikeda and Richardson, 1987). The higher half-saturation constant was commonly attributed to a particular involvement of GTP in transcript initiation. Apparent affinity constants for initial GTP binding (Martin and Coleman, 1989) and GTP consumption of elongating T7 RNAP (Guajardo and Sousa, 1997) were found to be 600  $\mu\text{M}$  and 41  $\mu\text{M}$ , respectively. An overall affinity constant for GTP should thus be located somewhere between the measured thresholds for pure initiation and pure elongation. The fact that the  $K_M$  value for GTP was shown to have a similar value in this study when compared with other substrate nucleotides suggests that the contribution of initiation to overall affinity toward GTP is smaller than previously believed. However, it is presently unclear whether our estimation of a lower  $K_{M,GTP}$  was related specifically to the template applied, or whether a generalization is possible to span other DNA sequences also. From the estimation of individual rate constants for initiation, elongation, and termination of the plasmids used, and from the relationships given for the derived model constants (cf. Table I), it is obvious that there is some variation to be expected among affinity constants for different vector systems.

Another aspect strongly influencing RNA synthesis rate is the type and concentration of ions present in solution (Chamberlin and Ring, 1973; Davis and Breckenridge, 1999; Gunderson et al., 1987; Ikeda and Richardson, 1987; Kern and Davis, 1997; Maslak and Martin, 1994) that is, T7 RNAP requires  $\text{Mg}^{2+}$  for the enzyme to develop its catalytic activity (Chamberlin and Ring, 1973). The effective concentration of freely dissolved ions, however, is reduced by a portion being bound to reaction components, such as, for example, nucleotides and nucleic acids (Record et al., 1976; Storer and Cornish-Bowden, 1976). Salt concentrations for optimum transcription performance are thus found to vary greatly with the concentration of ion-binding reactants

(Davis and Breckenridge, 1999; Ikeda and Richardson, 1987; Maslak and Martin, 1994).

When only total ion concentrations added to a reaction system are given, a comparison of results among different studies is problematic. A distinction between metal ions freely dissolved in solution, rather than of total salts added is desirable. Free ion concentrations were thus computed in this study for the reaction conditions prevalent at the start and termination of each experiment. No significant change was noted for the free  $Mg^{2+}$  concentration over the course of the reaction, with a calculated maximum deviation of 8% under our applied conditions. However, the choice of reaction conditions may well affect ionic strength significantly over the process duration. This was observed, for instance, by Young et al. (1997), who noted a drop in the level of free  $Mg^{2+}$  of several millimolar with time of RNA production due to the high removal of magnesium through precipitated  $Mg_2PPi$ . This removal of magnesium was due to the high initial NTP concentration applied and the resultant dynamics of, in particular, NTP concentration and  $Mg_2PPi$  during nucleotide conversion. An a priori neglect of the dynamics of ion concentrations is thus not advisable. It is conceivable to consider the time dependence of ionic strength when the algorithm for calculation of ion species concentrations is used simultaneously with numerical integration.

The presented transcription model for the T7 RNA polymerase mechanism is, in principle, equipped to be integrated as a functional subunit into an overall model for protein biosynthesis, which is part of ongoing research. For the representation of such systems, the mechanism of RNA degradation and its impact on ribosomal translation need to be considered in addition to RNA synthesis rate, together with the kinetics involved within the ribosomal apparatus. Such a procedure for defining appropriate modular building-blocks and assuring their compatibility pursues the ambitious goal of designing functional subunits of simultaneous transcription/translation with the aid of modeling and simulation of nonlinear cellular dynamics. In this context, models allowing full use of nucleotide sequence information available in genomic databases will provide new possibilities in prediction capacity.

The authors thank all members of this consortium, represented by Prof. V. Erdmann (Institute of Biochemistry, FU Berlin, Germany), Prof. A. Spirin (Institute for Protein Research, Pushchino, Russia), Dr. G. Stadler (Institute for Bioanalytics, Göttingen, Germany), and Dr. A. Roeder (Roche Molecular Diagnostics, Penzberg, Germany) for fruitful discussions.

## APPENDIX

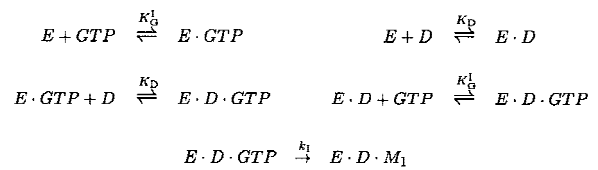
### Reaction Scheme

The elementary reaction steps underlying the assumed transcription mechanism catalyzed by bacteriophage T7 RNAP are shown in Figure A1. The reaction scheme is divided into four sections: transcription initiation; elongation; competi-

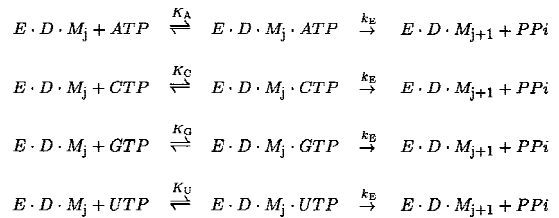
tive inhibition; and termination. Equilibrium constants,  $K_i$ , have been used for reversible reaction steps. Irreversible reactions are characterized by the rate constants  $k_i$ ,  $k_E$ , and  $k_T$ , respectively.

(1) *Initiation*. Although GTP is commonly agreed upon to be the initial nucleotide integrated in RNA synthesis by T7 RNAP, there exists some controversy about the order of substrate binding at transcript initiation. GTP was proposed to bind first and invoke promoter specificity (Basu and Maitra, 1986; Sen and Dasgupta, 1993). Other studies have suggested promoter bond formation to occur prior to nucleotide association (Jia and Patel, 1997; Martin and Coleman, 1987). We have thus chosen a mechanism allowing for random-order binding of both initial GTP and T7 promoter. Gunderson et al. (1987) found the binding constants of T7 RNAP for promoter association to be  $1.0 \cdot 10^8 M^{-1}$  and for nonpromoter binding to be  $2.1 \cdot 10^4 M^{-1}$ . Due to this observed high specificity of T7 RNAP to its promoter, nonspecific bond formation with the nonpromoter sequence was neglected in this study. An irreversible reaction step completes the initiation reaction.

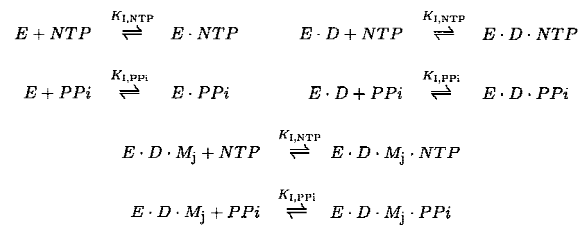
#### Initiation



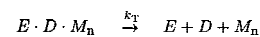
#### Elongation



#### Competitive Inhibition



#### Termination



**Figure A1.** Reaction scheme of transcription by T7 RNA polymerase.  $E$  denotes the biocatalyst T7 RNA polymerase;  $D$  is the promoter on the DNA template;  $M_j$  stands for an RNA transcript of length  $j$ ;  $E \cdot D \cdot M_j$  are intermediate enzymatic complexes containing the enzyme itself, DNA, and the nascent RNA. NTP denotes all nucleoside triphosphates, which are competing with the correct substrate nucleotide. PPi represents inorganic pyrophosphate.

(2) *Elongation.* Association of nucleotides to the ternary complex of T7 RNAP, DNA, and RNA (transcription complex denoted  $E \cdot D \cdot M_j$  in Fig. A1) is assumed to be independent of nucleotide sequence of the RNA chain. Transcription translocation occurs in an irreversible reaction step.

(3) *Competitive inhibition.* Competing nucleotides are able to bind to free enzyme, to T7 RNAP is bound to its promoter, and to the transcription complex. Erroneous incorporation of nucleotide into the growing RNA chain is ignored due to its low probability of occurrence, which has been reported to be in the order of magnitude of  $10^{-5}$  (Blank et al., 1986). Taking into account the adverse effect of inorganic pyrophosphate on RNA synthesis rate (Cunningham and Ofengand, 1990; Guajardo and Sousa, 1997), the nucleotide binding site of both freely dissolved and elongating T7 RNAP can be temporarily occupied by inorganic pyrophosphate.

(4) *Termination.* After completion of RNA synthesis, the ternary transcription complex is spontaneously disintegrated in a single reaction step releasing the RNA product.

## Termination Efficiency

Termination efficiency (TE) of transcription at a particular termination site,  $j$ , can be defined as the relative ratio of the concentration of  $RNA_j$  corresponding to this site with respect to the totality of RNA products synthesized from this DNA template (von Hippel and Yager, 1991). In mathematical form, this definition translates into:

$$TE_j = \frac{C_{RNA,j}}{\sum_{i=1}^R C_{RNA,i}} \quad (A1)$$

Alternatively, termination efficiency may be expressed in terms of rate of formation of the RNA products occurring in Eq. (A1); that is:

$$TE_j = \frac{V_j}{\sum_{i=1}^R V_i} \quad (A2)$$

Under rate-limiting conditions, any of the applicable rates given by Eqs. (3)–(7) may be substituted into Eq. (A2), in order to compute the efficiency of transcription termination.

For simplification, we assume in the following that  $V_i = V_{max,i}$  (for  $i = 1$  to  $R$ ) at excess substrate levels, and calculate the termination efficiency at the example of terminator,  $T_\phi$ , contained on plasmid pEThyd. Using Eqs. (11) and (12), it then follows for Eq. (A2) that:

$$TE_\phi = \frac{k_{eff,\phi}}{k_{eff,\phi} + k_{eff,pET1} + k_{eff,bla}} = \left( 1 + \frac{\left( \frac{1}{k_{I,\phi10}} + \frac{n_{T,\phi} - 1}{k_E} + \frac{1}{k_{T,\phi}} \right)}{\left( \frac{1}{k_{I,\phi10}} + \frac{n_{T,pET1} - 1}{k_E} + \frac{1}{k_{T,pET1}} \right) + \left( \frac{1}{k_{I,\phi10}} + \frac{n_{T,bla} - 1}{k_E} + \frac{1}{k_{T,bla}} \right)} \right)^{-1} \quad (A3)$$

Substituting the parameter values listed in Table III and the respective RNA lengths from Table II, we estimate  $TE_\phi = 0.67$ . Given the simplifying assumption of maximum transcript formation rates for either RNA product synthesized from pEThyd, this value compares quite favorably with the termination efficiency of 0.80 reported elsewhere for this terminator (Lyakhov et al., 1998). Eq. (A3) may be generalized to form:

$$TE_j = \frac{1}{1 + x \left( \frac{1}{k_I} + \frac{n_{T,j} - 1}{k_E} + \frac{1}{k_{T,j}} \right)} \quad (A4)$$

with the lumped parameter  $x$  given by:

$$x = \sum_{\substack{i=1 \\ i \neq j}}^R \left( \frac{1}{k_I} + \frac{n_{T,i} - 1}{k_E} + \frac{1}{k_{T,i}} \right)^{-1} \quad (A5)$$

From close inspection of Eq. (A4), it can generally be concluded that, with increasing value of termination rate constant  $k_{T,j}$ , the efficiency of transcription termination specific to site  $j$  is found to be increased also, whereas small  $k_{T,j}$  values lead to a reduced termination efficiency.

## NOMENCLATURE

bp	base pair
BSA	bovine serum albumin
$C_D$	total promoter concentration ( $\mu M$ )
$C_{E,t}$	total concentration of T7 RNAP ( $\mu M$ )
$C_{RNA}$	total concentration of RNA ( $\mu M$ )
$C_{RNA,i}$	concentration of RNA species $i$ ( $\mu M$ )
$C_{NTP,j}$	concentration of nucleoside triphosphate $j$ ( $\mu M$ )
$C_{PPi}$	concentration of inorganic pyrophosphate ( $\mu M$ )
dpm	disintegrations per minute
DTT	dithiothreitol
$f_{j,i}$	relative portion of base $j$ contained in RNA $i$ (%)
$k_d$	first-order rate constant for enzyme inactivation ( $\text{min}^{-1}$ )
$k_E$	rate constant for elongation ( $\text{s}^{-1}$ )
$k_{eff}$	overall rate constant for transcription ( $\text{s}^{-1}$ )
$k_I$	rate constant for initiation ( $\text{s}^{-1}$ )
$K_i$	respective equilibrium constant ( $M$ )
$K_I$	inhibition constant for respective metabolite ( $\mu M$ )
$K_M$	Michaelis–Menten constant for respective substrate ( $\mu M$ )
$k_T$	rate constant for termination ( $\text{s}^{-1}$ )
$m_i$	ratio of RNA species $i$ to total measured RNA (g/g)
$n_i$	transcript length for RNA species $i$ (kb)
$N$	number of ribonucleic bases
NTP	nucleoside 5'-triphosphate
PAGE	polyacrylamide gel electrophoresis
PPi	inorganic pyrophosphate
$R$	number of RNA species synthesized from a given DNA template
RNAP	RNA polymerase
$t$	time (min)
Tris	Tris(hydroxymethyl)aminomethane
$V$	transcription rate ( $\mu M/\text{min}$ )
$V_{max}$	maximum transcription rate ( $\mu M/\text{min}$ )
$X$	measured radioactivity (dpm/ $\mu L$ )

### Greek letters

$\alpha$	reduced scintillation measurement due to sample immobilization (%)
----------	--

$\gamma$  specific activity of radionuclide (Bq/mmol)  
 $\delta$  dilution factor of nonlabeled and [ $^3\text{H}$ ]-labeled UTP

## References

- Bailey JE. 1998. Mathematical modeling and analysis in biochemical engineering: Past accomplishments and future opportunities. *Biotechnol Prog* 14:8–20.
- Basu S, Maitra U. 1986. Specific binding of monomeric bacteriophage T3 and T7 RNA polymerases to their respective cognate promoters requires the initiating ribonucleoside triphosphate (GTP). *J Mol Biol* 190:425–437.
- Blank A, Gallant JA, Burgess RR, Loeb LA. 1986. An RNA polymerase mutant with reduced accuracy of chain elongation. *Biochemistry* 25: 5920–5928.
- Brendel V, Trifonov EN. 1984. A computer algorithm for testing potential prokaryotic terminators. *Nucl Acids Res* 12:4411–4427.
- Carrier TA, Keasling JD. 1997. Mechanistic modeling of prokaryotic mRNA decay. *J Theor Biol* 189:195–209.
- Cha S. 1968. A simple method for derivation of rate equations for enzyme-catalyzed reactions under the rapid equilibrium assumption or combined assumptions of equilibrium and steady state. *J Biol Chem* 243: 820–825.
- Chamberlin M, Ring J. 1973. Characterization of T7-specific ribonucleic acid polymerase I. General properties of the enzymatic reaction and the template specificity of the enzyme. *J Biol Chem* 248:2235–2244.
- Cleland WW. 1963. The kinetics of enzyme-catalyzed reactions with two or more substrates or products I. Nomenclature and rate equations. *Biochim Biophys Acta* 67:104–137.
- Cunningham PR, Ofengand J. 1990. Use of inorganic pyrophosphatase to improve the yield of *in vitro* transcription reactions catalyzed by T7 RNA-polymerase. *Biotechniques* 9:713–714.
- Davis RH, Breckenridge NC. 1999. Modeling of repeated-batch transcription for production of RNA. *J Biochem* 71:25–37.
- Draper NR, Smith H. 1981. Applied regression analysis. New York: John Wiley & Sons.
- Golomb M, Chamberlin M. 1974. Characterization of T7-specific ribonucleic acid polymerase. *J Biol Chem* 249:2858–2863.
- Good NE, Winget GD, Winter W, Connelly TN, Izawa S, Singh RMM. 1966. Hydrogen ion buffers for biological research. *Biochemistry* 5: 467–477.
- Guajardo R, Lopez P, Dreyfus M, Sousa R. 1998. NTP concentration effects on initial transcription by T7 RNAP indicate that translocation occurs through passive sliding and reveal that divergent promoters have distinct NTP concentration requirements for productive initiation. *J Mol Biol* 281:777–792.
- Guajardo R, Sousa R. 1997. A model for the mechanism of polymerase translocation. *J Mol Biol* 265:8–19.
- Gunderson SI, Chapman KA, Burgess RR. 1987. Interactions of T7 RNA polymerase with T7 late promoters measured by footprinting with methidiumpropyl-EDTA-iron(II). *Biochemistry* 26:1539–1546.
- Ikeda RA, Lin AC, Clarke J. 1992. Initiation of transcription by T7 RNA polymerase at its natural promoters. *J Biol Chem* 267:2640–2649.
- Ikeda RA, Richardson CC. 1987. Enzymatic properties of a proteolytically nicked RNA polymerase of bacteriophage T7. *J Biol Chem* 262: 3790–3799.
- Jermutus L, Ryabova LA, Plückthun A. 1998. Recent advances in producing and selecting functional proteins by using cell-free translation. *Curr Opin Biotechnol* 9:534–548.
- Jia Y, Patel SS. 1997. Kinetic mechanism of GTP binding and RNA synthesis during transcription initiation by bacteriophage T7 RNA polymerase. *J Biol Chem* 272:30147–30153.
- Job D, Soulié JM, Job C. 1988. Potential memory and hysteretic effects in transcription. *J Theor Biol* 134:273–289.
- Käpylä J, Hyytiä T, Lahti R, Goldman A, Baykov AA, Cooperman BS. 1995. Effect of D97E substitution on the kinetic and thermodynamic properties of *Escherichia coli* inorganic pyrophosphatase. *Biochemistry* 34:792–800.
- Kern JA, Davis RH. 1997. Application of solution equilibrium analysis to *in vitro* RNA transcription. *Biotechnol Prog* 13:747–756.
- Kern JA, Davis RH. 1999. Application of a fed-batch system to produce RNA by *in vitro* transcription. *Biotechnol Prog* 15:174–184.
- Kramer K, Hock B. 1996. Recombinant single-chain antibodies against *s*-triazines. *Food Agric Immunol* 8:97–109.
- Kukko-Kalske E, Lintunen M, Inen MK, Lahti R, Heinonen J. 1989. Intracellular P<sub>i</sub> concentration is not directly dependent on amount of inorganic pyrophosphatase in *Escherichia coli* K-12 cells. *J Bacteriol* 171:4498–4500.
- Langer RS, Hamilton BK, Colton CK. 1977. Enzymatic regeneration of ATP. II. Equilibrium studies with acetate kinase and adenylate kinase. *AIChE J* 23:1–10.
- Lyakhov DL, He B, Zhang X, Studier FW, Dunn JJ, McAllister WT. 1998. Pausing and termination by bacteriophage T7 RNA polymerase. *J Mol Biol* 280:201–213.
- Martin CT, Coleman JE. 1987. Kinetic analysis of T7 RNA polymerase–promoter interactions with small synthetic promoters. *Biochemistry* 26:2690–2696.
- Martin CT, Coleman JE. 1989. T7 RNA polymerase does not interact with the 5'-phosphate of the initiating nucleotide. *Biochemistry* 28: 2760–2762.
- Martin CT, Muller DK, Coleman JE. 1988. Processivity in the early stages of transcription by T7 RNA polymerase. *Biochemistry* 27:5755–5762.
- Maslak M, Martin CT. 1994. Effects of solution conditions on the steady-state kinetics of initiation of transcription by T7 RNA polymerase. *Biochemistry* 33:6918–6924.
- Mauch K, Arnold S, Posten C, Reuss M. 1997. Computer algebra systems in model-building and model-analysis for bioprocesses. 15th IMACS World Congr 2:171–178.
- May O, Habenicht A, Mattes R, Syltatk C, Siemann M. 1998. Molecular evolution of hydantoinases. *Biol Chem* 379:743–747.
- Menninger JR. 1983. Computer simulation of ribosome editing. *J Mol Biol* 171:383–399.
- Merkel W, Schwarz A, Fritz S, Reuss M, Krauth K. 1996. New strategies for estimating kinetic parameters in anaerobic wastewater treatment plants. *Wat Sci Tech* 34:393–401.
- O'Sullivan WJ, Perrin DD. 1964. The stability constants of metal–adenine nucleotide complexes. *Biochemistry* 3:18–26.
- Pozhitkov AE, Lavrik IN, Sergeev MM, Kochetkov SN. 1998. Kinetic analysis of reaction catalyzed by the phage T7 RNA polymerase. *Mol Biol* 32:78–82.
- Record MT Jr, Lohman TM, de Haseth P. 1976. Ion effects on ligand–nucleic acid interactions. *J Mol Biol* 107:145–158.
- Rhodes G, Chamberlin MJ. 1974. Ribonucleic acid chain elongation by *Escherichia coli* ribonucleic acid polymerase. *J Biol Chem* 249: 6675–6683.
- Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Sen R, Dasgupta D. 1993. Interaction of ribonucleotides with T7 RNA polymerase: Probable role of GTP in transcription initiation. *Biochem Biophys Res Commun* 195:616–622.
- Smith RM, Alberty RA. 1956. The apparent stability constants of ionic complexes of various adenosine phosphates with divalent cations. *J Am Chem Soc* 78:2376–2380.
- Sousa R, Patra D, Lafer EM. 1992. Model for the mechanism of bacteriophage T7 RNAP transcription initiation and termination. *J Mol Biol* 224:319–334.
- Stahl SJ, Zinn K. 1981. Nucleotide sequence of the cloned gene for bacteriophage T7 RNA polymerase. *J Mol Biol* 148:481–485.
- Stiege W, Erdmann VA. 1995. The potentials of the *in vitro* protein biosynthesis system. *J Biotechnol* 41:81–90.
- Storer AC, Cornish-Bowden A. 1976. Concentration of mgATP<sup>2-</sup> and

- other ions in solution. Calculation of the true concentrations of species present in mixtures of associating ions. *Biochem J* 159:1–5.
- Studier FW, Moffat BA. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* 189:113–130.
- Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Meth Enzymol* 185:60–89.
- Újvári A, Martin CT. 1996. Thermodynamic and kinetic measurements of promoter binding by T7 RNA polymerase. *Biochemistry* 35: 14574–14582.
- von Hippel PH, Yager TD. 1991. Transcript elongation and termination are competitive kinetic processes. *Proc Natl Acad Sci USA* 88:2307–2311.
- Young JS, Ramirez WF, Davis RH. 1997. Modeling and optimization of a batch process for *in vitro* RNA production. *Biotechnol Bioeng* 56: 210–220.