In Vitro RNA SYNTHESIS SHOULD BE COUPLED TO PYROPHOSPHATE HYDROLYSIS

Leonard Peller

Cardiovascular Research Institute
University of California
San Francisco, California 94143

Received February 10, 1975

Summary: Thermodynamics limits the number average degree of polymerization of RNA transcribed *in vitro* from DNA without hydrolysis of the pyrophosphate product to substantially less than 100 nucleotide units in size. The existence of considerably longer chains is initially compensated by unreacted nucleoside triphosphates and later by the presence of short chains of dubious biological utility. Pyrophosphate hydrolysis raises the upper limit of this average size by a factor of the order of 10^3 and effectively defers the approach to equilibrium.

In vitro studies of RNA transcription from a DNA template are made *without* coupling the synthesis to the strongly exergonic hydrolytic conversion of the pyrophosphate product to orthophosphate - a condition strikingly different from the situation *in vivo*. The baleful consequences of the omission of inorganic pyrophosphatase, the agent catalyzing this hydrolysis, from systems both for DNA and RNA synthesis have gone unappreciated. Two related difficulties attend nucleic synthesis in these incomplete *in vitro* systems: a) inability to achieve sufficiently high degrees of polymerization (1) and b) failure to avoid a broad distribution of chain sizes brought about by pyrophosphorolysis (2). Here we stress the first deficiency in connection with RNA synthesis.

Two basic observations applying to RNA transcription combined with a third general proposition governing linear polymerizations are central to the argument. (I) *All RNA transcripts are initiated by a ribonucleoside triphosphate*. There are no special initiator molecules. That initiation may be preferentially by purine nucleosides (3) is not germane here. (II) *The synthetic product is single stranded and is ultimately dissociated from the DNA template* (Fig. 1). The DNA is in effect a catalyst making certain nucleotide sequences available for transcription but remains unaltered at the conclusion of the reaction - in contradistinction to the replication process.

(III) The number average degree of polymerization at any time in the course of the reaction must be less than the final equilibrium value. This restriction does not apply to weight and higher average degrees of polymerization nor does it hold for any averages involving the polymeric species excluding the monomers (vide infra). Their magnitudes will depend on certain characteristics of the kinetic steps.

The equilibrium constant for formation of 3',5'-phosphodiester linkages can be defined by the following mass action expression

$$\frac{[3',5']_{eq}[PP]_{eq}}{[5'PPP]_{eq}[3'OH]_{eq}} = K \qquad (1)$$

where $[3',5']_{eq}$ is the equilibrium molar concentration of these bonds, $[PP]_{eq}$ that of pyrophosphate while $[5'PPP]_{eq}$ and $[3'OH]_{eq}$ are the concentrations of reactive groups. At all times $[5'PPP] = [3'OH] \equiv n$ where n is the concentration of chains including nucleoside triphosphate monomers. If the initial concentration of the latter is m_o, then $[3',5] = [PP] = \sum_{i=1}^{\infty} (i-1)m_i = \sum_{i=1}^{\infty} im_i - \sum_{i=1}^{\infty} m_i = m_o - n$ where m_i is the molar concentration of chains with i-1 bonds and degree of polymerization i. When such a closed system attains equilibrium

$$\frac{[m_o - n_{eq}]^2}{[n_{eq}]^2} = K \qquad (2)$$

whence

$$\langle i \rangle_{eq} \equiv \frac{m_o}{n_{eq}} = 1 + \sqrt{K} \qquad (3)$$

The mode of derivation of this simple result (obtained in a different context and by a more labored analysis some years ago (4)) illustrates the validity of proposition III that $\langle i \rangle \leq \langle i \rangle_{eq}$. If this were not so, the relative concentration of bonds formed could transiently exceed that at equilibrium.

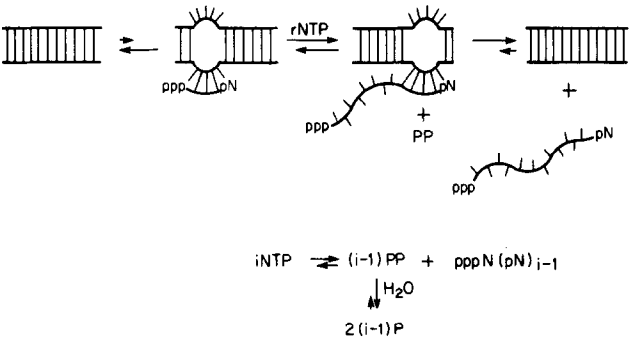

Figure 1. Three stages in the synthesis of RNA from a DNA template with the i mer transcript finally desorbed. Bottom: a summary of the steps with the pyrophosphate product converted to orthophosphate. The relative lengths of the arrows for the forward and reverse reaction qualitatively convey the degree of reversibility.

Taking as an outrageous upper limit $K = 10^4$, we find that $\langle i \rangle_{eq} \simeq 100$ which corresponds to a molecular weight of 33,000 daltons and represents a species size of the order of tRNA. This is well below that of mRNA and could code for polypeptides of molecular weights of only 3×10^3. The above value for K is absurdly high with better than half the free energy of hydrolysis of ATP to AMP and PP being made available for the polymerization. More realistic estimates have K ranging from 10 to 100 (1, 2).

RNA chains as large as 2.5×10^6 daltons have been reported as arising from *in vitro* transcription of *inter alia* T7 DNA (5). The presence of chains much longer than $\langle i \rangle_{eq}$ doesn't violate thermodynamic strictures provided that there is an abundance of shorter chains not rendered insoluble in the customary trichloracetic acid precipitation of synthesized material (3). At the earliest stages of this polymerization the unreacted ribonucleoside triphosphates will assure that this equilibrium average is not exceeded. With the passage of time the newly synthesized chains must be of considerably smaller size than those initially transcribed in conformity with the thermodynamic limit. (The equilibrium value for the number average degree of polymerization of species commencing with dinucleoside phosphates is $2 + \sqrt{K}$ but this does not represent an inviolable upper limit - only a dismal ultimate desti-

nation). It cannot be emphasized too strongly that the appearance of large molecules in this system is a consequence of the significantly slower rate of initiation than of propagation of chains.

The in vivo synthesis of RNA plainly doesn't occur in a closed system as ribonucleoside triphosphate substrates are constantly being regenerated and the RNA produced is gathered into the protein synthesizing machinery. Of the intracellular elements missing from RNA syntheses in vitro the easiest to supply is the enzyme inorganic pyrophosphatase.

Conversion of pyrophosphate to two orthophosphate ions is so nearly complete that $[PP]_{eq} \simeq \sqrt{\frac{4m_o}{K_h}}$ where K_h is the equilibrium constant for the hydrolysis. Then we obtain as the equilibrium upper limit for <i> an earlier result (1)

$$\langle i \rangle_{eq} \simeq \sqrt{\frac{KK_h}{4m_o}} \qquad (4)$$

With a small value of K = 10 and $K_h = 10^5$ (1, 6) and $m_o \simeq 10^{-4}$M, $\langle i \rangle_{eq} \simeq 5 \times 10^4$ or a molecular weight of 1.7×10^7 daltons - an order of magnitude larger than mRNA species.

The RNA arising from selective DNA transcription must inevitably display a heterogeneity in size as comparisons of in vitro with in vivo products reveal (7). Pyrophosphorolysis must act to widen further the natural size distribution - a characteristic of all depolymerization reactions. The hydrolysis of pyrophosphate also serves to extend the time scale of the approach to an unwanted broad equilibrium size distribution by a factor of $K_h/4m_o \simeq 10^9$ (2). Both the inherent irreversibility of this conversion (large K_h) and the dilution of the system (2 orthophosphate ions appear for every phosphodiester bond as in the bottom of Fig. 1) combine to diminish the importance of the depolymerization process under these circumstances.

ACKNOWLEDGEMENTS: The author is grateful to Professor E. Peter Geiduschek for helpful critical comments. Supported by NHLI grant HL-06285.

References

1. Peller, L. (1966) Proc. Nat. Acad. 55, 1025.
2. Peller, L. submitted for publication.
3. Chamberlin, M. (1974) The Enzymes, 3rd Ed., Vol. X, Boyer, P.D., Ed., chapter 10, Academic Press, New York.
4. Peller, L. and Barnett, L. (1962) J. Phys. Chem. 66, 680.
5. Dunn, J.J. and Studier, F.W. (1973) Proc. Nat. Acad. Sci. 70, 1559.
6. Alberty, R.A. (1969) J. Biol. Chem. 244, 3290.
7. Brody, E. and Geiduschek, E.P. (1970) Biochemistry 9, 1300.