Lawrence Berkeley National Laboratory

Recent Work

Title

SEQUENCES AND EFFICIENCIES OF PROPOSED mRNA TERMINATORS

Permalink

<https://escholarship.org/uc/item/22q2k79t>

Author

Tinoco Jr., Ignacio

Publication Date

1977-11-01

Submitted to Nature

RECEIVED
LIBRARY
BIOCHEMICAL LABORATORY

LBL-7239
Preprint *c. 2*

APR 27 1978

LIBRARY AND
DOCUMENTS SECTION

SEQUENCES AND EFFICIENCIES OF
PROPOSED mRNA TERMINATORS

James E. McMahon and Ignacio Tinoco, Jr.

November 1977

Prepared for the U. S. Department of Energy
under Contract W-7405-ENG-48

TWO-WEEK LOAN COPY

*This is a Library Circulating Copy
which may be borrowed for two weeks.
For a personal retention copy, call
Tech. Info. Division, Ext. 5716*

LBL-7239

c. 2

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

SEQUENCES AND EFFICIENCIES OF PROPOSED mRNA TERMINATORS

The base sequences at the 3' end of different messenger RNA molecules show a pattern of a GC rich region followed by five to eight U's.^{1,2} Gilbert's hypothesis¹ that the stability of the template duplex DNA (or of the mRNA-DNA hybrid, or both) determines termination of transcription, implies that the DNA termination sequences, and thus the mRNA sequences at the 3' end, are not unique. Rather, the DNA sequence of the template presumably expresses a certain structural stability, and this pattern of stabilities is responsible for the termination of transcription. Using these ideas we have found five (but only five) sequences in ϕ X174 DNA which have a common pattern of stabilities with previous terminator sequences. These sequences are proposed to be responsible for termination of transcription in ϕ X174. Three of the five sequences have five or six A's following a GC region instead of the T's present in known terminators. The lengths of messenger RNAs predicted on this basis are in excellent agreement with published observations. However, sequence determination of the messengers will be necessary to rigorously test our proposal.

We have analyzed those termination sequences given by Gilbert in terms of free energy of formation of each of the Watson-Crick nearest-neighbor sequences. There are some data available from thermodynamic studies of synthetic deoxypolynucleotides,³ but we have used the complete set of free energy values determined for RNA double strands.⁴ The relative stabilities are

For simplicity in representing the sequence in these terms, the nearest-neighbors are labeled from 1 to 6 in order of decreasing stability. Note that these numbers characterize a sequence of two base pairs, not a single base pair. For example, GC represents the sequence

The known termination sequences^{1,2} are characterized by a most stable nearest-neighbor (GG) followed by two variable terms, then a series of weak nearest-neighbors such that at least five of seven of the subsequent terms are 6 on the stability scale. Known termination sequences contain oligo U sequences, but we propose that oligo A sequences are also possible; it is the stability of the oligo dA·dT region which is important.

An analysis of the sequence of the plus strand of ϕ X174 DNA⁵ reveals only five locations where this pattern of nearest-neighbor stabilities occurs (Fig. 1). There are three promoters^{6,7} in vitro and the exact location of these has been determined by Sanger.⁵ Using these three known mRNA initiation sites and our five proposed termination sequences, the expected mRNAs may be listed according to their molecular weights and which protein codes they contain. A comparison between this prediction and the messenger RNA found in vivo⁸ is given in Table 1. In general excellent agreement is found.

Hayashi et al.⁸ separated fourteen peaks of different molecular weight mRNAs by electrophoresis. By hybridization of these fractions with restriction fragments, those authors were able to orient nine mRNA fractions on the ϕ X174 genome. All nine are accounted for by predicted mRNAs; seven show exact agreement in both size and content. (Fractions 0 and IX are less certain.)

Hayashi et al.⁸ reported that RNAs starting at promoter A' are rapidly degraded in vivo, with the possible exception of RNA species I. In Table 1, the predicted RNAs which do not have counterparts in vivo are those that initiate at promoter A'; and the observed fractions in vivo which do not correspond to predicted messengers are of smaller molecular weight and thus may represent degradation products.

Of particular significance is the observation that for each of four predicted terminators, one or more observed mRNA fractions may be assigned. No additional terminators need be proposed to explain the observed data. None of the predicted mRNAs which are terminated by T₁ are observed experimentally, therefore, they are not shown in Table 1. Since T₁ differs from the other predicted terminators by the presence of a guanine residue interrupting the AT-rich region and thus contributing two stronger stability terms, it may be expected to be a less efficient terminator. However, Axelrod⁷ has observed a rho-dependent terminator in vitro which maps at a location in agreement with our predicted T₁.

Since Hayashi and co-workers⁸ have given values for the relative number of molecules of each of the fractions mapped, the efficiencies of each predicted terminator may be calculated. This calculation is performed using only those fractions assigned to starts at promoter A. The results in Table 2 are consistent with Axelrod's in vitro observation that T₄ is a strong terminator while T₂ is less effective. In addition, the terminator designated T_H is seen to have a similar efficiency to T₂, while T_G is even less effective.

The relative efficiencies of the terminators may be related to their primary sequence. The terminator sequences are listed in order of decreasing efficiency in Fig. 1. All terminators require a strong

nearest-neighbor interaction outlined in the box at the left. The other boxed region is composed of ten nucleotides. The efficiency of termination appears correlated with the number of adjacent T or A residues and their distance from the strong nearest-neighbor interaction. Thus T_6 starting in the first position outlined is more efficient than A_6 in the same place or T_6 starting farther to the right. All of these are more efficient than A_5 starting in the first position. Given our assumptions about the importance of weak nearest-neighbor interactions in this region, it seems reasonable to speculate that the presence of G after A_5 in T_1 may make this a much weaker, or totally ineffective, terminator. Yet, in view of the in vitro evidence this may act effectively in the presence of rho factor. There is obviously much still to be learned about terminator efficiency and sequence, but we stress the importance of duplex stability (DNA·DNA and DNA·RNA) here.

On the basis of terminator efficiencies calculated from promoter A one may calculate the relative number of molecules for each terminator for another promoter, given the assignment of any one mRNA for that promoter. Thus for promoter G, the distribution of RNAs may be calculated given the assignment of RNA XIII to T_2 . Good agreement is found for fraction 0 and III, but one finds that fraction VI contains nearly twice as many molecules as expected. This may be explained by reference to Table 1, where it is seen that two mRNAs of the same molecular weight, but arising from different promoters and terminators, have probably been collected together as fraction VI.

Determination of the base sequences of the mRNAs can verify the proposed terminators.

This work was supported in part by National Institutes of Health

Grant GM 10840 and by the Division of Biomedical and Environmental
Research of the United States Energy Research and Development Admin-
istration.

James E. McMahon

Ignacio Tinoco, Jr.

Department of Chemistry and Chemical
Biodynamics Laboratory
University of California
Berkeley, California 94720

Terminator*	Position†	Sequence and Stability Terms	
T ₄	3961	G G A G G C T T T T T T A T G G T T	
		1 4 4 1 2 4 6 6 6 6 6 5 5 4 1 4 6	
T _H	2635	C G A C C C T A A A T T T T T T T C C	
		3 4 4 1 1 4 5 6 6 5 6 6 6 6 6 4 1	
T ₂	854	A A A G G T A A A A A A C G T T C T	
		6 6 4 1 4 5 6 6 6 6 6 4 3 4 6 4 4	
T _G	2308	A G C G G T A A A A A T T T T A A T	
		4 2 3 1 4 5 6 6 6 6 5 6 6 6 5 6 5	
T ₁	4429	T A C C C C A A A A A G A A A G G T	
		5 4 1 1 1 4 6 6 6 6 4 4 6 6 4 1 4	

Fig. 1 Proposed terminator sequences and stability terms for ØX174. The approximate location of the last residue transcribed is before the right-hand solid line, except for terminator T_H (indicated by the dashed line).

* Notation from reference 7; T_G and T_H refer to terminators labeled "start of gene G" and "start of gene H," respectively.

† Indicates location of first base listed, using numbering of reference 5.

Table 1. mRNAs: Comparison with experiment

<i>in vivo</i> *			Prediction		
Fraction	MW x 10 ⁻⁶	Proteins	MW x 10 ⁻⁶	Proteins	Promoter/ Terminator [†]
			2.61	ABCDEJFGHABCDE	P _{A'} / T ₂
0	2.3	(C) DEFGHABCDE	2.29	BCDEJFGHABCDE	P _A / T ₂
			2.00	DEJFGHABCDE	P _G / T ₂
I	1.8	ABCDEFGH	1.83	ABCDEJFGH	P _{A'} / T ₄
II	1.5	BCDEFGH	1.52	BCDEJFGH	P _A / T ₄
			1.38	ABCDEJF(G)	P _{A'} / T _H
			1.27	ABCDEJF	P _{A'} / T _G
III	1.3	(C) DEFGH	1.23	DEJFGH	P _G / T ₄
IV	1.1	BCDEFG	1.07	BCDEJF(G)	P _A / T _H
V	.98	BCDEF	.96	BCDEJF	P _A / T _G
			.79	DEJF(G)	P _G / T _H
VI	.82	(C) DEF	.78	ABCDE	P _{A'} / T ₂
VII	.73	?			
VIII	.66	?	.67	DEJF	P _G / T _G
IX	.6	BCDE	.46	BCDE	P _A / T ₂
X	.52	?			
XI	.43	?			
XII	.32	?			
XIII	.23	D(E)	.17	DE	P _G / T ₂

* From reference 8

[†] Notation as in reference 7

TABLE 2. EFFICIENCY OF TERMINATORS

<u>Promoter</u>	<u>Terminator</u>	<u>Fraction</u>	<u>Relative No. of Molecules</u>	<u>Terminator Efficiency</u>	<u>Sequence</u>
P _A	T ₂	0	0.3	-	
P _A	T ₄	II	2.2	79	GGCTTTTTTATGG
P _A	T _H	IV	3.1	53	CCCTAAATTTTTT
P _A	T ₂	IX	9.5	50	GGTAAAAACGTT
P _A	T _G	V	3.0	34	GGTAAAAATTTTA

Relative No. of Molecules
Observed Predicted

P _G	T ₂	0	0.3	0.3
P _G	T ₄	III	2.3	2.1
P _G	T _H	VI	5.2	2.9
P _G	T ₂	XIII	9.0	(9.0)
P _G	T _G	-	-	2.8

References

1. Gilbert, W. in RNA Polymerase (ed. Losick, R. & Chamberlin, M.) 193-205 (Cold Spring Harbor Laboratory, New York, 1976).
2. Subramanian, K. N., Ghosh, P. K., Dhar, R., Thimmapaya, B., Zain, S. G., Pan, J. & Weissman, S. M. Prog. in Nucleic Acid Res. and Mol. Biol. 19, 157-164 (1976).
3. Wells, R. D., Larson, J. E., Grant, R. C., Shortle, B. E., & Cantor, C. R. J. Mol. Biol. 54, 465-497 (1970).
4. Borer, P. N., Dengler, B., Tinoco, I., Jr. & Uhlenbeck, O. C. J. Mol. Biol. 86, 843-853 (1974).
5. Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A. III, Slocombe, P. M. & Smith, M. Nature 265, 687-695 (1977).
6. Smith, L. H., & Sinsheimer, R. L. J. Mol. Biol. 103, 699-710 (1976).
7. Axelrod, N. J. Mol. Biol. 108, 753-770 & 771-779 (1976).
8. Hayashi, M., Fujimura, F. K., & Hayashi, M. Proc. Natl. Acad. Sci., USA 73, 3519-3523 (1976).

This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720