Model for regulation of the histidine operon of Salmonella
(gene regulation/antitermination control/leader peptide/alternative stem model/mRNA structure)

H. Mark Johnston*, Wayne M. Barnes*, Forrest G. Chumley*, Lionello Bossi†, and John R. Roth*

*Department of Biology, University of Utah, Salt Lake City, Utah 84112, and †Department of Biochemistry, Washington University Medical School, St. Louis, Missouri 63110

Communicated by Bruce N. Ames, September 4, 1979

ABSTRACT A model is proposed that accounts for regulation of the histidine operon by a mechanism involving alternative configurations of mRNA secondary structure (the alternative stem model). New evidence for the model includes sequence data on three regulatory mutations. The first (his01242) is a mutation that deletes sequences needed to form the attenuator mRNA stem and causes constitutive operon expression. The second mutation (his09654) is a His's ochre (UAA) mutation in the leader peptide gene; the existence of this mutation constitutes evidence that the leader peptide gene is translated. The third mutation (his09663) is remarkable. It neither generates a nonsense codon nor affects a translated sequence; yet, it is suppressible by amber suppressors. We believe this mutation causes a His's phenotype by interfering with mRNA secondary structure. The suppressibility of the mutation is probably due to disruption of the attenuator stem by ribosomes that read through the terminator codon of the leader peptide gene. This explanation is supported by the observation of derepression of a wild-type control region in the presence of an amber suppressor. Evidence is presented that hisT mutants (which lack pseudouridine in the anticodon arm of histidine tRNA) may cause derepression of the hist operon by slowing protein synthesis in the leader peptide gene.

The histidine operon is a cluster of nine genes whose expression increases in response to histidine starvation. Regulation seems to be achieved without any purely regulatory protein. Here we present a model for regulation of this operon and some preliminary data that support that model.

The background
Our current understanding of this regulatory mechanism rests on several sorts of data. A large number of constitutive mutants have been isolated. These mutants fall into six classes, one mapping near the operon (hisO), the others (hisR, S, T, U, and W) mapping at separate positions far from the operon (1, 2). None of the unlinked mutations appears to affect a repressor protein; all seem to affect the amount or the structure of histidyl-tRNA (3-9). Therefore, the mechanism regulating operon expression must sense the level of histidyl-tRNA (10, 11). Five of the six classes of "regulatory" mutations generate a derepression signal without directly affecting the regulatory region. The hisO mutations may be the only class that directly affects the regulatory apparatus. The hisO mutations are dominant and affect only genes contiguous (cis) to the mutant site; mutations in the other regulatory genes are recessive (refs. 12 and 13; unpublished results).

Transcription of his operon DNA in vitro revealed that the control region includes a barrier to transcription (14). This barrier, termed the attenuator site, is apparently removed or damaged by hisO constitutive mutations. It was suggested that regulation is achieved by altering the frequency with which transcription crosses this attenuator site.

By use of an in vitro system that permits transcription of the operon either coupled with or uncoupled from translation (15), it was demonstrated that transcription through the attenuator and into the structural genes of the operon occurs only when translation is occurring simultaneously.

The DNA sequence of the histidine operon control region (hisO) was determined by Barnes (16). This sequence includes two particularly interesting features: (i) A sequence with dyad symmetry is present which, if transcribed into message, would permit formation of a perfect 14-base-pair stem and loop. This mRNA stem includes a region rich in G and C residues and is followed by nine U residues. Structures of this type have been associated with message-termination signals in several other systems (17-23). It seemed likely that the attenuator site (14) might encode this mRNA structure (the attenuator stem). (ii) The control region was also found to contain a sequence that could encode a peptide of 16 amino acids. The gene for this peptide includes seven adjacent histidine codons! It was suggested that the translation requirement for in vitro operon expression might involve this tiny gene. The run of histidine codons could provide a sensitive means for determining the concentration of histidyl-tRNA to which operon control responds.

The role of the leader peptide gene in regulation has been speculative. No direct evidence has previously linked this gene or the putative attenuator stem to operon control. Here we propose a model mechanism for regulation of the histidine operon and present preliminary data that support several aspects of this model.

Model for regulation of histidine operon
The model assumes that formation of the 14-base-pair stem (the attenuator stem) in the histidine mRNA causes RNA polymerase to terminate transcription at the run of U residues following that stem. If formation of this stem is prevented, then RNA polymerase proceeds across the attenuator site and into the structural genes. Formation of the attenuator stem is prevented when the promoter-proximal (5') half of the stem is already involved in alternative secondary structure at the time that the distal (3') portion of the stem is synthesized. Fig. 1 presents three possible configurations that might be assumed by the leader portion of the histidine mRNA. Fig. 1a and c presents structures that include the attenuator stem (EF) and, thus, would cause repression of operon expression. The configuration in Fig. 1b does not include the attenuator stem and, thus, would permit full operon expression. Fig. 1c shows three main stem-loop structures which are designated AB, CD, and EF (attenuator). These letters refer to stretches of message se-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.
The repressed operon is presented in Fig. 1a. It is assumed that when an excess of histidyl-tRNA is available, the first ribosome follows RNA polymerase closely through the leader sequence designated in Fig. 1. The configuration in Fig. 1b is generated by an alternative pairing arrangement of these same stretches to form new stem-loop structures BC and DE. We assume that these stems, once formed, are relatively stable on the time scale involved in making the regulatory decision.

The model proposes that the configuration assumed by the message depends on whether sequences are available for pairing with newly made mRNA. For example, newly synthesized sequence C (see Fig. 1) can form a BC stem if B is not already paired, if B is involved in a preexisting AB stem, C will be unable to pair until the subsequent D sequence is made and the CD stem can form.

The model assumes that availability of mRNA sequences for pairing is affected by the position of the first ribosome on the nascent mRNA. It is presumed that ribosomes are capable of disrupting mRNA secondary structure as they translate the message. More explicitly, we assume that a ribosome disrupts 12 bases of message structure ahead (at the 3' side) of the codon occupying the aminoacyl-tRNA site on the ribosome (26). The mechanism proposed works even if the extent of mRNA structure disruption is varied slightly from this assumed value of 12.

The ramifications of these assumptions are presented in Fig. 1.

The Operon © 1980 Cold Spring Harbor Laboratory Press 0-87969-133-6
For conditions see www.cshpress.com/copyright.
of histidine codons would cause formation of the attenuator stem, (Fig. 1c) and lead to transcription termination. The fact that the ochre mutation is His" supports our model. Suppression of this mutation despite the low efficiency of ochre suppressors is consistent with the fact that very low operon expression (≤ 1/10 of basal levels) is sufficient for a His" phenotype.

A Novel Suppressor/Regulatory Mutation. A problem arose in accounting for some of the other suppressible His" mutants. The sequence of the leader peptide gene includes no codons that can be converted to UAG or to UGA by a single base substitution. Yet eight amber-suppressible mutations and one UGA-suppressible mutation have been found promoter-proximal to the attenuator. Mutation hisO9663 is one of these "amber" mutations; it causes the mRNA change seen in Fig. 2. It is apparent that this "amber" mutation neither generates a nonsense codon nor is it even located in a region thought to be translated, yet it causes a His" phenotype and is corrected by amber suppressors.

The behavior of this unusual mutation is understandable in terms of the alternative stem model presented above. The position of the base substitution is marked with an arrow (base 1) in Fig. 1. The base change alters the B sequence (Fig. 2), which destabilizes the BC stem (ΔG = -10.2 kcal reduced to ΔG = -5.8 kcal). We suggest that this destabilization prevents sufficient BC stem formation and causes CD and EF (Fig. 1a) to form under all circumstances, this blocks operon expression and leads to a His" phenotype. The suppressibility of this mutation is explained in Fig. 2. Suppression of the termination codon (UAG) of the leader peptide would allow the first ribosome to read out of the peptide gene to a UGA codon at the base of the attenuator stem (Fig. 2). A ribosome that follows the polymerase closely up to this point should disrupt the attenuator stem and thus permit polymerase to read into the structural genes. Since very little expression of the his operon (≤ 1/10 of fully repressed levels) is needed for a His" phenotype, relatively few ribosomes would need to progress to this point. Thus, we think that mutation hisO9663 owes its histidine requirement to excessive attenuator function. This interpretation is supported by the fact that hisO9663 is suppressed by the attenuator deletion hisO1242. It seems likely that the suppressibility of hisO9663 by amber suppressors is due to extension of the leader peptide. If this is true, these amber suppressors might affect expression of a wild-type operon.

Operon Derepression by Amber Suppressors. To test the effect of leader peptide extension on operon expression, we used an F" plasmid carrying a wild-type his control region fused to the lac operon of E. coli (unpublished data). The β-galactosidase gene on this plasmid is expressed and regulated by the his promoter and control mechanism. This plasmid was transferred to a series of isogenic Salmonella strains carrying various nonsense suppressors, β-galactosidase levels were then assayed. Results are presented in Table 1.

Most of the suppressors tested caused an increase in operon expression. Observation of this effect depends on growing cells in rich medium. Although the influence of growth conditions is not yet understood, we propose that rich medium is needed if the ribosome is to keep pace with RNA polymerase while traversing the region of stems. If it falls behind (perhaps due to slight shortages of any of a variety of charged tRNAs), the attenuator stem can again form and cause repression.

Comparison of E. coli and Salmonella Leader Sequences. The DNA sequence of the his control region has been determined for E. coli (32) and Salmonella (18). Only two differences are seen among the 151 base pairs from the leader peptide AUG codon through the 5' base of the attenuator stem. This high degree of sequence homology contrasts with the general extent of sequence divergence between the two species, which has been estimated as 13% (33). The close homology suggests that almost every point of the sequence has been under selective pressure. The only two differences between the E. coli and Salmonella sequences in this region (bases 2 and 3 in Fig. 1) affect residues that are not base paired in either mRNA ar-

The Operon © 1980 Cold Spring Harbor Laboratory Press 0-87969-133-6
For conditions see www.cshpress.com/copyright.
Isogenic strains containing an F' plasmid on which the his operon control elements are fused to the lac genes of E. coli were assayed for β-galactosidase enzyme activity. The results presented are the average for two independent experiments. The β-galactosidase activity of the sup + strain averaged 10 units, and this was defined as a derepression level of 1.0. Cells growing exponentially in an amino acid-rich medium were assayed according to Miller (28). The suppressors used were characterized by Winston et al. (29). The broth in which the cells were grown contained Difco Bactryptone (10 g/liter) and NaCl (8 g/liter), with the following amino acids added (per liter): histidine (16 mg), serine (420 mg), glutamine (730 mg), tyrosine (18 mg), tryptophan (21 mg), phenylalanine (50 mg), and cysteine (36 mg). The full genotype of the strains, omitting the sup genotype indicated, is his-100 leu-114 tyr-16:Ts5/F'P000-1 lacI475/Tn10 his+G^+D^+ (his-lacZ) lac^*4T*. The nomenclature for these Tn10 and Tn5 insertions has been described (30, 31).

Table 1. Effect of nonsense suppressors on expression of his operon

<table>
<thead>
<tr>
<th>Suppressor genotype</th>
<th>his operon derepression level</th>
</tr>
</thead>
<tbody>
<tr>
<td>sup**</td>
<td>1.0</td>
</tr>
<tr>
<td>supE4a (Ser)</td>
<td>1.2</td>
</tr>
<tr>
<td>supE6a (Gln)</td>
<td>1.3</td>
</tr>
<tr>
<td>supF2b (Tyr)</td>
<td>3.1</td>
</tr>
<tr>
<td>supF4a (Leu)</td>
<td>3.1</td>
</tr>
<tr>
<td>supG6a (Lys)</td>
<td>2.2</td>
</tr>
<tr>
<td>supC6a (Tyr)</td>
<td>1.1</td>
</tr>
</tbody>
</table>

The model proposes that message termination, and therefore regulation of the his operon, is mediated through translation of the leader peptide gene. This was also suggested by Barnes (16). Suppressibility of an ochre mutation in this small gene demonstrates that the leader peptide gene is in fact translated. This explains the requirement for translation for in vitro operon expression (15) and strengthens our belief that the ribosome is the major positive factor regulating his operon expression. No purely regulatory proteins are involved in the model as presented above, but it remains possible that some protein or polypeptides serve to stabilize loop structures of many operons. It is unlikely that the his gene product modulates regulation in any direct way (38).

The control region includes His" mutations that are unstable and revert frequently to His+. We believe that all these mutations will, like hisO9653 and hisO9654, prove to affect the formation of mRNA stems or translation of the leader peptide. Their instability may be due to the large number of ways of correcting such stem sequences. Any secondary mutation that affects formation or stability of the attenuator stem would be expected to suppress these His" regulatory mutations. At the far left of the map is a group of stable His" mutations. These, we believe, will prove to remove or damage the his promoter.

Under maximally repressed conditions, the his operon is still expressed at a basal level. The details of how this basal level is maintained are not yet clear. The basal level does not merely reflect the extent to which the attenuator is unable to prevent readthrough. The hisO9663 and hisO9654 mutations owe their His" phenotype to attenuator function. Therefore, the attenuator stem appears capable of blocking virtually all transcription. We believe that the basal level will be explained by statistical fluctuations in ribosome position under repressing conditions. If, occasionally, the first ribosome is late in initiating transcription, or is slowed by fluctuations in concentration of any of the charged tRNAs, the attenuator stem would not form. The basal level, we expect, will reflect the frequency with which such events occur under repressing conditions.

Note Added in Proof: An alternative stem model similar to that described here has recently been devised for regulation of the trp operon of Escherichia coli (39).

Discussion

A model has been described proposing that alternative mRNA configurations are involved in regulation of the his operon. This alternative stem model is similar in principle to one proposed for the trp operon by Lee and Yanofsky (37). An important aspect of the model presented here is the assumption that a particular stem can prevent formation of a later alternative stem even if the second stem is energetically favored. This is possible if the time involved to shift from the first stem to the second is so great that RNA polymerase would pass the critical point for termination before the attenuator stem could form.

The model proposes that message termination, and therefore regulation of the his operon, is mediated through translation of the leader peptide gene. This was also suggested by Barnes (16). Suppressibility of an ochre mutation in this small gene demonstrates that the leader peptide gene is in fact translated. This explains the requirement for translation for in vitro operon expression (15) and strengthens our belief that the ribosome is the major positive factor regulating his operon expression. No purely regulatory proteins are involved in the model as presented above, but it remains possible that some protein or polypeptides serve to stabilize loop structures of many operons. It is unlikely that the his gene product modulates regulation in any direct way (38).

The control region includes His" mutations that are unstable and revert frequently to His+. We believe that all these mutations will, like hisO9653 and hisO9654, prove to affect the formation of mRNA stems or translation of the leader peptide. Their instability may be due to the large number of ways of correcting such stem sequences. Any secondary mutation that affects formation or stability of the attenuator stem would be expected to suppress these His" regulatory mutations. At the far left of the map is a group of stable His" mutations. These, we believe, will prove to remove or damage the his promoter.

Under maximally repressed conditions, the his operon is still expressed at a basal level. The details of how this basal level is maintained are not yet clear. The basal level does not merely reflect the extent to which the attenuator is unable to prevent readthrough. The hisO9663 and hisO9654 mutations owe their His" phenotype to attenuator function. Therefore, the attenuator stem appears capable of blocking virtually all transcription. We believe that the basal level will be explained by statistical fluctuations in ribosome position under repressing conditions. If, occasionally, the first ribosome is late in initiating leader peptide synthesis or is slowed by fluctuations in concentration of any of the charged tRNAs, the attenuator stem would not form. The basal level, we expect, will reflect the frequency with which such events occur under repressing conditions.

Note Added in Proof: An alternative stem model similar to that described here has recently been devised for regulation of the trp operon of Escherichia coli (39).
This work was supported by National Institutes of Health Grants GM23408 (to J.R.) and GM24956 (to W.B.). W.B. is the recipient of American Cancer Society Junior Faculty Research Award No. 1.