Proc. Natl. Acad. Sci. USA
Vol. 75, No. 9, pp. 4281–4285, September 1978
Biochemistry

DNA sequence from the histidine operon control region: Seven histidine codons in a row

(genetic regulation/transcription/attenuation)

WAYNE M. BARNES

Department of Biological Chemistry, Division of Biology and Biomedical Sciences, Washington University School of Medicine, St. Louis, Missouri 63110

Communicated by Bruce N. Ames, July 10, 1978

ABSTRACT **The DNA sequence of 250 base pairs preceding the first structural gene of the histidine operon of *Salmonella typhimurium* was determined by the dideoxy chain-termination method. Single-stranded DNA template was provided by an M13-histidine transducing phage constructed for the purpose by *in vitro* recombination. The termination site for the histidine leader RNA is identified by analogy with the *trp* operon leader termination sequence, and is 47 nucleotides before the start codon of the first structural gene *G*. Beginning 150 nucleotides before the end of the presumed leader RNA is a possible short protein-coding region with *seven* histidine codons in a row. It is proposed that the major mechanism of histodine operon control must involve a ribosome arrested at this run of histidine codons when histidine is limiting.**

Control of transcription of the histidine operon of *Salmonella typhimurium* and of the tryptophan operon of *Escherichia coli* involves both a variably efficient start site, the promoter, and a variably efficient stop site, the attenuator, both in the DNA genetic control region in front of the first structural gene (1–3). The small RNA that is made between the promoter and the attenuator, the "leader RNA," has been isolated and the sequence for the *trp* operon has been determined (4), but the analogous RNA from the *his* operon has not been so characterized.

Various large and small control molecules, sensing the metabolic situation of the cell *vis-a-vis* the need for histidine or tryptophan, mediate the action of RNA polymerase at these dual DNA control sites. The major control for the *trp* operon is exerted at the start of transcription by the *trp* repressor (5). The histidine operon has no repressor and, thus, its major control is exerted at the attenuator. Two control molecules have been implicated as accessories in controlling the histidine attenuator. Histidyl-$tRNA_{His}$ is strongly involved, its concentration being sensitively and inversely related to the level of expression of the operon (6). Although the product of the first structural gene *G* of the operon has been implicated in genetic control (7), its direct role has been questioned by the observation that the histidine operon of a *G* deletion mutant is controlled normally (8). Control at the histidine operon attenuator seems to require translation of something (2) and to involve some positive (antiterminating) factor (1, 2).

To help elucidate the actual control mechanism used by the histidine operon, I have determined the DNA sequence of the 250 base pairs immediately preceding the first structural gene *G*. Features of the DNA sequence suggest a direct role for the ribosome as a positive factor in control of transcription of the histidine operon.

MATERIALS AND METHODS

M13Hol 67 and 76 are single-stranded, recombinant transducing phages carrying the *His OGD* region of *S. typhimurium* in opposite orientations. They were grown under P2 conditions of physical containment, with F^- *E. coli* K-12 as host. Phage were precipitated from the culture supernatant with polyethylene glycol (9) and banded in CsCl, and the DNA was deproteinized with phenol. Primer DNA restriction fragments were identified and isolated as described (10,11) from the mini ColE1-His OGD plasmid pWB91 (10). [α^{32}P]dATP, specific activity of 100–150 Ci/mmol, was obtained from ICN. *E. coli* DNA polymerase I large fragment (12, 13) was from Boehringer Mannheim. Restriction enzyme *Hha* I was a gift from R. Roberts; *Hae* III was isolated according to Roberts *et al.* (14).

DNA Sequencing. The method of Sanger *et al.* (15) was applied with slight modifications. All enzymatic reactions were carried out in buffer A (50 mM NaCl/10 mM Tris·HCl, pH 7.9/10 mM $MgCl_2$/1 mM dithiothreitol). Mixtures of dNTPs and analogs [dideoxynucleoside triphosphates (ddNTPs) or arabinonucleoside triphosphates; P. L. Biochemicals] were prepared in water and stored frozen at 10 times the final concentrations shown in Table 1.

DNA primer fragment and M13Hol DNA single-strand template (1 pmol each) were denatured and reannealed in 50 μl of 10/8 strength buffer A, added to 0.1 pmol of dried [α-^{32}P]dATP, and divided into five aliquots of 8.5 μl. One microliter of the appropriate dNTP/analog mixture in water and 0.5 μl of DNA polymerase large fragment (0.5 unit) were added and the reactions were incubated in a 37° oven for 15 min. The dideoxy reactions were then chased with 1 μl of all four dNTPs at 0.5 mM, and the araC reaction with 1 μl of A, T, and G only at 1 mM, for a further incubation of 15 min at 37°. The reactions were terminated by addition of 10 μl of 20 mM EDTA and evacuated to dryness. Each DNA mixture was resuspended in 10 or 15 μl of 98% formamide/0.1% xylene cyanol/0.1% bromphenol blue/10 mM EDTA and heated to 90° for 30 sec; 4 or 5 μl was applied per gel sample.

Table 1. Mixtures of dNTPs and analogs

Final	ddA/dA	ddT/dT	ddG/dG	ddC/dC	araC/dC
Analog/dNTP	50	50	50	16	800
μM analog	100	100	100	33	1600
μM dNTP	0*	2	2	2	2

The other dNTPs were at 10 μM, except there was no dATP in the stock solutions.

* dATP was added separately, usually α-^{32}P-labeled, as appropriate, at a final concentration of 2 μM.

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.

Abbreviations: dN, deoxynucleotide; ddN, 2′,3′-dideoxynucleotide; A, adenosine mononucleotide or triphosphate, as appropriate, and similarly for G, guanosine; C, cytosine; araC, -T, or -U, arabinocytidine -thymidine, or -uracil.

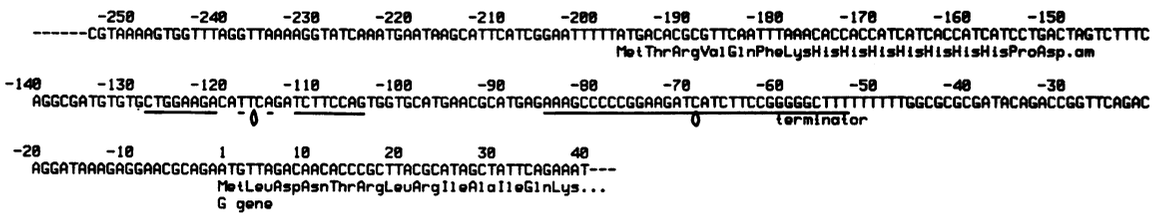

FIG. 1. DNA sequence from the histidine operon control region. Only one strand is shown. Centers of symmetry are indicated by ◊, and the symmetrical positions are underlined.

The sequence-analyzing gels were the thin(20 cm × 40 cm × 0.4 mm) 8% acrylamide/7 M urea gels described by Sanger and Coulson (16). Spacer material (clear rigid vinyl) was 0.015 inches (0.38 mm) thick, and sample wells were 5 mm wide. Fixing and autoradiography were as described (11, 17). Exposure times were 7–70 hr. Sequence data were handled by computer (18).

RESULTS

DNA Sequence. The DNA sequence of the 250 base pairs preceding the *hisG* gene of the histidine operon is presented in Fig. 1. Essentially all of the data for the DNA sequence in Fig. 1 were derived from the two experiments (repeated several times) shown in Fig. 2. The chain-termination sequencing method of Sanger *et al.* (15) was used, with *Hha* I or *Hae* III restriction fragments as primers and transducing phage M13Hol DNA as the single-stranded template. Fig. 2 *A* and *B* shows the result of a priming by fragment RH51 annealed to M13Hol 76 DNA (priming to the right on the map in Fig. 3). Enzyme recleavage by *Hha* I was used in this experiment, hence the cutoff at a molecular length of 220 nucleotides, the next *Hha* site. These gels are very easily and unambiguously interpreted. Note that sequence-specific variations in band intensity are similar to those found with the partial ribosubstitution method (11), i.e., cC_n (positions −63, −81, −108, and −174) and Aa_n (positions −179 and −184), where the lowercase letters represent the fainter bands. The araC channel is not particularly reliable nor, therefore, usually as useful as it is in Fig. 2*A*.

Some 50 nucleotides of sequence closer to the primer in this experiment have been run off the gel in Fig. 2*A*. These sequence data were previously provided by the partial ribosubstitution method (figure 5 in ref. 11) and have been confirmed by a shorter electrophoresis of the chain-termination experiment of Fig. 2 *A* and *B*. At the top end of the sequence in Fig. 2 *A* and *B*, the length of the T_9 run cannot be determined (although nine Ts can be counted), due to overexposure by the heavy band at the next *Hha* I restriction site (actually a double site at position −45). This problem was overcome by inserting a single ribonucleotide, rC, before incorporating the label (11, 15), and carrying out final recleavage with piperidine (11). The resulting bands were somewhat fuzzier than the bands when a portion of the same material was recleaved with *Hha* I, so only 10–12 additional nucleotides could be resolved beyond the run T_9. Nevertheless, this experiment provided 27 nucleotides of overlap with the sequence determined in Fig. 2*C*.

Fig. 2*C* presents a priming in the opposite direction (M13Hol 67 DNA as template) from a point 44 nucleotides into the *G* gene (*Hae* III fragment RZ54 as primer). This gel has two problems in interpretation. The first problem is that the dark bands all have an artifact shadow band under them. This would result if 5–10% of the primer molecules have one less base at their 5′ end. This heterogeneity could be due to exonuclease contamination in the restriction enzyme used to prepare the fragment. Whatever the cause of the shadow bands, they must be subtracted from the analysis using the variable intensity rules described above. For instance, since runs of As always appear as Aa_n, the first band in the patterns aAa (position +6), aA (position −15), and aAa_8 (position −25) must be considered an artifact. The sequence experiment in Fig. 2*C* has been repeated with longer electrophoresis time and with priming by another restriction fragment farther away from this sequence, *Hha* I fragment RH54 (not shown). In this experiment the artifact bands were not present, and the same sequence was indicated.

The other problem with interpretation of the data in Fig. 2*C* occurs at the positions 17–18. These gels are not completely denaturing and, at some sequences, strong secondary structure increases the mobility of the DNA strand being analyzed. This

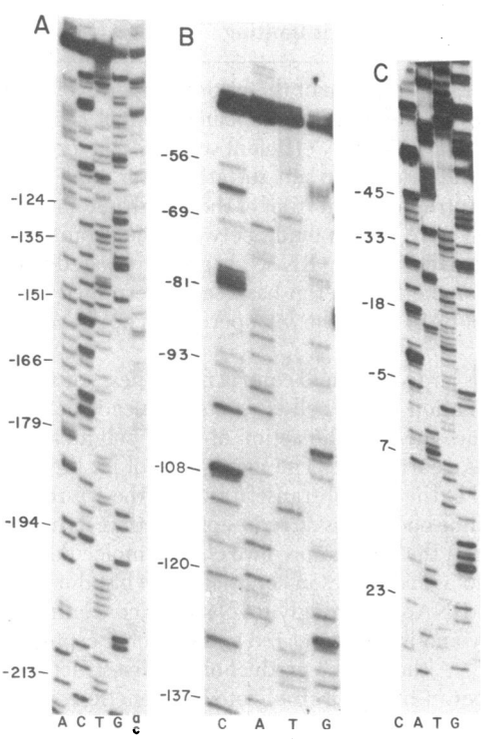

FIG. 2. Sequence data. These representative autoradiographs, generated by the chain-termination method (15, 16), present most of the primary data for the sequence of 250 base pairs presented in this paper. ab, an arabino CTP (araCTP) channel. The other channels used dideoxytriphosphates. (*A* and *B*) Sequence was generated by priming with *Hha* fragment RH51 (see map in Fig. 3) annealed to template M13Hol 76 DNA, with a longer electrophoresis time for the experiment in *B*. These sequencing experiments used two polymerase extension steps, a labeling extension and an analog-incorporation extension, separated by Sephadex G100 gel filtration, and followed by restriction enzyme recleavage (19, 11). The labeling extension was done according to the "limiting substrate extension" protocol (11). The second extension was done by the normal analog incorporation procedure described in *Materials and Methods*, except unlabeled dATP was used. After the analog incorporation (15-min incubation), 1 unit of restriction enzyme *Hha* was added to cleave at the restriction site at the edge of the primer. (*C*) Single extension was used, priming with *Hae* III fragment RZ54 on M13Hol 67 DNA. *B* is reproduced on a different scale, 4/7 actual size. Numbers on the left correspond to the position numbers in Fig. 1. Upper part of the gels is not shown.

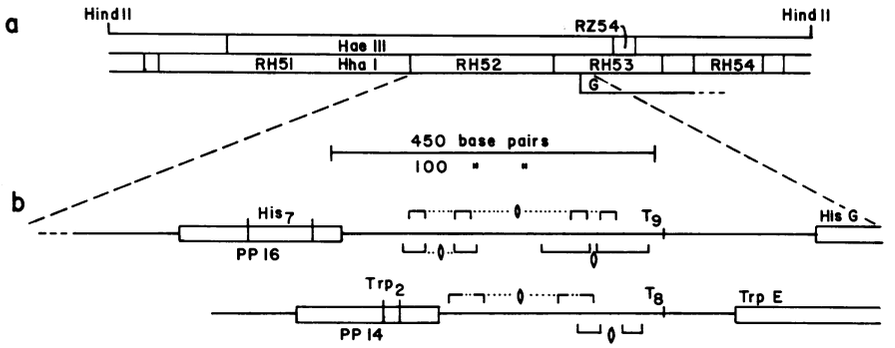

FIG. 3. Partial restriction map and summary of DNA sequence features. (*a*) Restriction sites of enzymes used to determine the histidine operon DNA sequence, *Hha* I and *Hae* III, are indicated. Some of the fragments are designated with their operational names, such as RZ54. (*b*) Features of the DNA sequence of the *Salmonella* histidine operon and the *E. coli* tryptophan operon (4, 21). The *Salmonella* tryptophan operon is not summarized here, but it is very similar to the *E. coli* tryptophan operon (22). The genetic control regions are drawn to scale, with the terminators aligned with each other. Centers of complete or hyphenated symmetry are denoted O, with the symmetrical sequences delimited by the brackets. RNA structures that form with symmetries denoted above each operon line are mutually exclusive with structures using symmetries denoted below each line. The boxed regions are proposed or actual peptide-coding regions. PP16, putative protein with 16 amino acids.

problem is known as a compression, which in this case is strong enough to be an inversion. The sequence at position 17 apparently reads . . . CGGGG . . . , with strang spacing in the apparent G_4 run. The amino acid sequence from this region is known (20), and it indicates that the true sequence must be . . . GCGGG . . . . The sequence thus established overlaps that determined in Fig. 2*B* near the run of nine A·T base pairs, which is part of the terminator signal of the attenuator. The length of this run can be easily counted in experiments on each DNA strand.

DISCUSSION

Features of DNA Sequence. Features of the DNA sequence are summarized in Fig. 3. The DNA sequence presented is bounded on the right by the coding region of the *His G* gene. This gene boundary was identified by comparison with the NH_2-terminal amino acid sequence of the G protein (20). We have recently determined the rest of the *G* gene DNA sequence and amino acid sequence (R. N. Husson and W. M. Barnes, unpublished data; D. Piszkiewicz, B. Tilley, T. Rand-Meir, and S. M. Parsons, unpublished data) and it is contiguous with the sequence presented here. The promoter will not be discussed here, since its location is not known, although it is probably just upstream of the sequence shown in Fig. 1.

Ultimate identification of the genetic control sites of the operon will require knowledge of the DNA sequence of mutants, the sequence of the leader RNA, and the binding sites of putative control proteins. In the current absence of any such data for the histidine operon, it is still possible to propose very probable identifications of various DNA sites important in control of the operon by inspection of the DNA sequence for *a priori* features and by analogy with other systems.

The easiest site to identify by analogy with other sequences is the terminator of the attenuator, i.e., the probable 3′ end of the leader RNA. The best comparison is with the leader RNA molecules of the mechanistically related *trp* operons of *E. coli* and *Salmonella* (21, 22), but there also exists valid similarity with other short RNA molecules transcribed by *E. coli* RNA polymerase, such as λ *oop* RNA (23), λ 6S RNA (24), and a small φ80 RNA (25). The common features at the 3′ end of these RNAs are a run of six to eight Us preceded by a high G-C region containing hyphenated (imperfect) symmetry. The histidine operon leader region also contains these features: There is a run of nine Ts ending 47 nucleotides before the *G* gene, and this T_9 is preceded by a region of *perfect* symmetry (centered at position −69) sixteen base pairs on each side containing *eight* G-C base pairs in a row. If these features are really termination signals, the histidine operon signals are the most extreme observed to date and, naively, it is perhaps surprising that RNA polymerase can ever read through this site. Biochemical measurements of the efficiency of this site indicate that termination *in vitro* is at least 8 efficient (1).

The center of symmetry at position −69 happens to contain two symmetrically related *Mbo* II recognition sites (GAAGA). These sites are included in a partial repeat of this sequence centered at position −127. Since the partially repeated sequence contains a center of symmetry, it can also be considered as an inverted repeat. RNA transcribed from these regions can be expected to form several alternative base-paired structures.

Two mutually exclusive sets of the RNA structures that could form as a result of the observed symmetry are indicated with brackets above or below the line representing the histidine operon in Fig. 3. Two mutually exclusive RNA structures for the *trp* operon leader, which have been proposed by Lee and Yanofsky (21), are similarly indicated for the *trp* operon in Fig. 3. They have proposed a model for attenuator control of the *trp* operon that is based in part on the importance and mutual exclusiveness of these RNA structures. If their proposal is valid for the histidine operon, the structures in question are even stronger and, as shown by the overlaps between upper and lower brackets in the diagrams, even more mutually exclusive.

Seven Histidine Codons in a Row. About 140 nucleotides before the putative terminator (positions −197 to −146), at what is probably the center of the leader RNA of the histidine operon, there is a potential peptide-coding region sixteen codons long that contains *seven* histidine codons in a row! This observation immediately suggests a simple model of genetic regulation in the attenuator region of the histidine operon. The speed of translation of seven histidine codons in a row should be highly sensitive to the level of histidyl-tRNA, and consequently the level of histidine, in the cell. I propose that a ribosome arrested before the seven histidine codons must have, by some unknown mechanism, an antitermination effect on RNA polymerase at the attenuator some 120 nucleotides away.*

Thus the positive effector that has been indicated for the histidine operon by previous *in vitro* and genetic analyses (1, 2, 26) is apparently a ribosome. This is consistent with every observation in the previous analyses, which have shown that the RNA polymerase terminates efficiently in the absence of translation (1, 2) and that complete translation machinery is

* After this observation was made, I was informed that in 1967 R. G. Martin and B. Ames proposed, but never published, a similar model involving histidyl-tRNA control of the operon by translational coupling through a series of histidine codons in a leader peptide.

necessary for derepression (2). Thus, the absence of any ribosome must be regarded as equivalent to complete, smooth translation of the seven-histidine leader peptide. We are left with a slow or arrested ribosome having an antiterminating effect on RNA polymerase, and no ribosome or a smoothly (and therefore briefly) translating ribosome having no effect on an otherwise efficient termination process. No role is proposed for the putative 16-amino-acid peptide itself.

Fig. 4 illustrates the molecular situation postulated for the moment of decision whether to transcribe through the terminator. What could be the mechanistic connection between an arrested ribosome and antitermination, particularly over such a distance? One possible model that comes to mind supposes that a termination factor such as rho (27) is necessary for termination, and this factor must enter upstream from the coding region of the nascent message, move along it, and catch up to RNA polymerase from behind to activate termination (28). If there is no entry site between the His_7 and the termination site, an arrested ribosome would then physically block this factor.

There are, however, three lines of evidence that rho factor itself is not involved at the histidine attenuator. First, rho factor is not required *in vitro* for termination at the poly(T) sequences of the short phage RNAs mentioned above, nor at the *trp* attenuator (29), nor at the histidine attenuator (1). Second, a survey of 10 rho (*suA*) mutants by Winkler (30) found no significant effect on the histidine attenuator *in vivo*, although such genetic experiments are complicated by the fact that rho is apparently such an important protein for the cell that absolute rho-defective mutants are not known. Some rho dependence has in fact been found for the *trp* attenuator *in vivo* (31). The third line of evidence that rho is not involved comes from comparative DNA sequence analysis of two terminators that are absolutely dependent on rho factor *in vitro*, the terminator t_R of λ (32), and the termination site in the gene for tyrosine tRNA (33). Neither of these terminators has the run of Ts, and each has either a G·C-rich region before the termination site *or* a region of symmetry, but not both. On the other hand, both of these rho-dependent terminators has the sequence -C-A-A-T-C-A-A- at the point of termination, while no sequence similar to this is found in the rho-independent small λ or φ80 RNAs, nor in the histidine control DNA. [The *trp* leader contains the sequence -C-A-A-T-C-A-G-, ending 28 nucleotides before the T_8. Perhaps this explains the rho effect observed for *trp* (31)]. These differences in features of the sequence at demonstrably (*in vitro*) rho-dependent termination sites and rho-independent termination sites suggest that there may be at least two classes of terminators that can be recognized from features of their nucleotide sequence alone. If so, then the terminator of the histidine attenuator is in the rho-independent class. Of course, it is still possible that some other termination factor is involved, but there is no evidence for it.

In the absence of any good evidence for a termination factor at the histidine attenuator, it is best to assume none is involved. How then could an arrested ribosome affect rho-independent termination? The mechanism of antitermination depends on the mechanism of rho-independent termination to be counteracted. RNA polymerase pauses at rho-catalyzed termination sites (32) and at sequences resembling rho-independent termination sites (34, 35). This pause may be as long as 1 min (32), and it may be caused by the difficulty of melting the DNA double helix at the high G·C region preceding the termination site (35). An unusually stable RNA·DNA hybrid in the high G·C region (36), the resulting displaced DNA loop (which may form a Gierer stem and stabilize the RNA·DNA hybrid complex; ref. 37 and R. G. Martin, personal communication), or a structure in the nascent RNA (21) may then somehow interact with the RNA polymerase to cause termination. A slight destabilization of the RNA structure might then counteract termination. It is implicit that the RNA structures must be on the edge of stability in order that they be responsive to any mechanism of genetic control. Therefore, a very small destabilization should be sufficient for a genetic control effect.

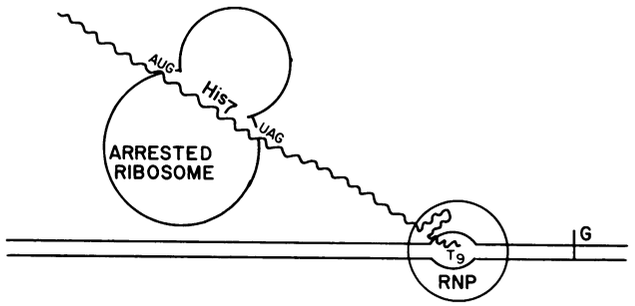

FIG. 4. Proposed molecular situation at the moment of the key histidine operon control decision: whether or not RNA polymerase should read through the termination signal of the attenuator. The situation depicted is that obtaining when the histidyl-tRNA is in low concentration due to limiting histidine. (Wavy line) Nascent leader RNA; (straight line) DNA; T_9, terminator. Coding sequence is shown for putative 16-amino-acid protein with seven histines in a row. One of the possible secondary structures of the leader RNA that may activate termination (21) is indicated within the domain of the RNA polymerase (RNP). Possible RNA·DNA hybrid which may be important at termination (36) is not shown.

It is possible that a ribosome arrested on the nascent leader RNA would exert a force like a rock on a string and pull out (destabilize) whichever RNA structure is important for termination. The source of this force would be water molecules imparting Brownian motion on this scale. Once the structure that causes RNA polymerase to pause is destabilized, RNA polymerase proceeds instantly into the structural genes of the operon, and the control decision has been made.

A similar small coding region, 14 codons long with two adjacent *trp* codons, has been found in the leader region of the tryptophan operon of *E. coli* (21) and *Salmonella* (22). For *E. coli*, this region has been shown to bind ribosomes at the initiator codon AUG (38) and, although the 14-amino-acid peptide has not yet been physically observed, efficient translation has been proved by the observation of fusion proteins in a *trp-lac* fusion (39) and in internal *trp* operon deletions (40). Based on these observations, Lee and Yanofsky (21) have proposed an arrested-ribosome model for attenuator control that is similar to the one proposed here for the histidine operon. In their model, however, complete translation of the short peptide is required to activate termination, in contrast to the model proposed here for histidine, in which the absence of a ribosome is equivalent to complete translation. Zurawski *et al.* (41) have found in the *E. coli pheA* leader region a coding sequence for a possible leader peptide with seven phenylalanines (not all contiguous).

Histidyl-tRNA and Control Regardless of the actual mechanism of antitermination by the ribosome, nearly all of the genetically and physiologically observed effects of histidyl-tRNA activity on expression can be explained by this arrested-ribosome model. One puzzle in interpretation of histidyl-tRNA effects in the past has been an observed hypersensitivity of expression of the histidine operon toward histidyl-tRNA levels: mutants *hisW*, *hisU*; and *hisR* decrease histidyl-tRNA levels by only 15–60%, yet they have 3- to 10-fold effects on expression (6). An arrested-ribosome mechanism is expected to be hypersensitive to slight changes in histidyl-tRNA activity, since any effect on speed of translation of a histidine codon would be multiplied by seven adjacent histidine codons.

The constitutive effect of $hisT^-$ tRNA, which lacks a pseudoracil modification (42), is at first glance harder to explain, however, since it is present in normal concentration (6), is

charged normally (6), and, since it is not lethal, apparently functions in protein synthesis. Nevertheless, since *hisT*$^-$ cells do grow somewhat more slowly than wild type (6), it is reasonable that the imperfect tRNA is in fact somewhat slow in translation, due perhaps to a slightly weaker or stronger interaction with the ribosome. Alternatively, *hisT*$^-$ tRNA in the ribosomal A site may interact unfavorably with an adjacent $tRNA_{His}$ is the P site, thus greatly slowing translation of *consecutive* histidine codons, while not having an appreciable effect in normal coding regions, where adjacent histidine codons are rare (suggestion of B. Ames and J. Roth). The expected hypersensitivity of the arrested ribosome model allows reconsideration of previous data indicating a role for histidyl-tRNA synthetase in control of the histidine operon (43). The apparent positive control effects observed can be explained by assuming that the synthetase is acting merely as a "sponge" to slightly reduce the availability of charged tRNA for translation.

Many interesting classes of mutations in the control region are predicted by the arrested ribosome model. Most of these would be "down" mutations and classified as promoter mutations (26 1). These down mutations include mutations in the ribosome-binding site for the leader peptide, creation of an ochre codon at position −179, and frameshift mutations between the AUG codon and the His_7. For instance, deletion of one base pair in this region would change the *his* codons to a run of threonine and isoleucine codons, and thus possibly put the operon under (experimentally testable) combined threonine and isoleucine control. Perhaps unfortunately for this prediction, such a mutation would extend the peptide coding region to a length of 33 codons (including 5 serine codons), so the increased time of translation might be compensatingly constitutive for the operon, and the mutation might not have a "down" phenotype. If the arrested ribosome model is right, the predictable classes of mutations almost certainly exist among the 100 control mutations already known (26), and the 50 odd promoter mutations recently isolated by Johnston and Roth (personal communication). It will be interesting to test the model by determining the DNA sequence of these mutations.

I thank F. Sanger, S. Nicklen, and A. Coulson for communicating their DNA sequencing methodology prior to publication; R. G. Martin, B. Ames, S. Artz, and J. Roth for discussions on the mechanism of histidine operon regulation; D. Kennell for criticizing the manuscript; and M. Scott and R. Wrenn for help with the computer. This work was supported in part by U.S. Public Health Service Grant 1 R01 GM24956-01.

1. Kasai, T. (1974) *Nature (London)* **249**, 523–527.
2. Artz, S. & Broach, J. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3452–3457.
3. Bertrand, K., Korn, L., Lee, F., Platt, T., Squires, C. L., Sqires, C. & Yanofsky, C. (1975) *Science* **189**, 22–26.
4. Squires, C., Lee, F., Bertrand, K., Squires, C. L., Bronson, M. J. & Yanofsky, C. (1976) *Proc. Natl. Acad. Sci. USA* **103**, 351–381.
5. Jackson, E. N. & Yanofsky, C. (1973) *J. Mol. Biol.* **76**, 89–99.
6. Lewis, J. A. & Ames, B. N. (1972) *J. Mol. Biol.* **66**, 131–142.
7. Meyers, M., Blasi, F., Bruni, C. B., Deeley, R. C., Kovach, J. S., Levinthal, M., Mullinix, K. P., Vogel, T. & Goldberger, R. F. (1975) *Nucleic Acids Res.* **2**, 2021–2036.
8. Scott, J. F., Roth, J. R. & Artz, S. W. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 5021–5025.
9. Yamamoto, K., Alberts, B., Benzinger, R., Lawhorne, L. & Treiber, G. (1970) *Virology* **40**, 734–744.
10. Barnes, W. M. (1977) *Science* **195**, 393–394.
11. Barnes, W. M. (1978) *J. Mol. Biol.* **119**, 83–99.
12. Setlow, P., Brutlag, D. & Kornberg, A. (1972) *J. Biol. Chem.* **247**, 224–231.
13. Klenow, H., Overgaard-Hansen, K. & Patkar, S. A. (1971) *Eur. J. Biochem.* **22**, 371–381.
14. Roberts, R. J., Breitmeyer, J. B., Tabachnik, N. F. & Myers, P. A. (1975) *J. Mol. Biol.* **91**, 121–123.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
16. Sanger, F. & Coulson, A. R. (1978) *FEBS Lett.* **87**, 107–110.
17. Air, G. M., Sanger, F. & Coulson, A. R. (1976) *J. Mol. Biol.* **108**, 519–533.
18. Staden, R. (1977) *Nucleic Acids Res.* **4**, 4037–4051.
19. Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* **94**, 441–448.
20. Piszkiewicz, D., Rand-Meir, T., Theodor, '. & Parsons, S. M. (1977) *Biochem. Biophys. Res. Commun.* **78**, 833–838.
21. Lee, F. & Yanofsky, C. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4365–4369.
22. Lee, F., Bertrand, K., Bennett, G. & Yanofsky, C. (1978) *J. Mol. Biol.* **121**, 193–217.
23. Dahlberg, J. E. & Blattner, F. R. (1973) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **32**, 664 (abstr.).
24. Lebowitz, P., Weissman, S. H. & Radding, C. M. (1971) *J. Biol. Chem.* **246**, 5120–5127.
25. Pieczenik, G., Barrell, B. G. & Gefter, M. L. (1972) *Arch. Biochem. Biophys.* **152**, 152–165.
26. Ely, B., Fankhauser, D. B. & Hartman, P. E. (1974) *Genetics* **78**, 607–631.
27. Roberts, J. (1969) *Nature (London)* **224**, 1168–1174.
28. Adhya, S., Gottesman, M., de Crombrugghe, B. & Court, D. (1976) in *RNA Polymerase* (Cold Spring Harbor Laboratories, Cold Spring Harbor, NY), pp. 719–730.
29. Lee, F., Squires, C. L., Squires, C. & Yanofsky, C. (1976) *J. Mol. Biol.* **103**, 383–393.
30. Winkler, M. E. (1978) *J. Bacteriol.* **135**, 721–725.
31. Korn, L. M. & Yanofsky, C. (1976) *J. Mol. Biol.* **103**, 395–409.
32. Rosenberg, M., Court, D., Shimatake, H., Brady, C. & Wulff, D. L. (1978) *Nature (London* **272**, 414–423.
33. Kupper, H., Sediya, T., Rosenberg, M., Egan, J. & Landy, A. (1978) *Nature (London* **272**, 423–428.
34. Maizels, N. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 3585–3589.
35. Gilbert, W. (1976) in *RNA Polymerase* (Cold Spring Harbor Laboratories, Cold Spring Harbor, NY), pp. 193–205.
36. Neff, N. F. & Chamberlin, M. J. (1978) *J. Biol. Chem.* **253**, 2455–2460.
37. Gierer, A. (1966) *Nature (London)* **212**, 1480–1481.
38. Platt, T., Squires, C. & Yanofsky, C. (1976) *J. Mol. Biol.* **103**, 411–420.
39. Schmeissner, U., Ganem, D. & Miller, J. H. (1977) *J. Mol. Biol.* **109**, 303–326.
40. Miozzari, G. F. & Yanofsky, C. (1978) *J. Bacteriol.* **133**, 1457–1466.
41. Zurawski, G., Brown, K., Killingly, D. & Yanofsky, C. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4271–4275.
42. Singer, C. E., Smith, G. R., Cortesi, R. & Ames, B. N. (1972) *Nature (London) New Biol.* **238**, 72–74.
43. Wyche, J. H., Ely, B., Cebula, T. A., Snead, M. C. & Hartman, P. E. (1974) *J. Bacteriol.* **117**, 708–716.