Nucleic Acids Research, 1993, Vol. 21, No. 25 5875–5881

Transcription of a silkworm $tRNA_c^{Ala}$ gene is directed by two AT-rich upstream sequence elements

Fakhruddin A.Palida[1], Charles Hale[1,2] and Karen U.Sprague[1,2,*]
[1]Institute of Molecular Biology and [2]Department of Biology, University of Oregon, Eugene, OR 97403, USA

Received September 20, 1993; Revised and Accepted November 19, 1993

ABSTRACT

A region within 35 nucleotides upstream of the transcription initiation site of a variety of silkworm Class III templates is absolutely required for transcription *in vitro*. To determine whether the activity of this region can be attributed to a particular sequence element, we systematically replaced 4–5 bp segments of the region upstream of a silkworm $tRNA_c^{Ala}$ gene. We show that replacement of either of two AT-rich blocks markedly impairs promoter function, whereas replacement of other sequences has little or no effect. Additional mutants were constructed to test whether base composition or sequence is important for function of the AT blocks. We find that some sequences are more effective than others, but that various AT-rich sequences can direct transcription at a high level. Possible mechanisms by which such elements could act are discussed.

INTRODUCTION

tRNA promoters are generally considered to consist of two highly conserved elements, the A and B boxes, which are located within the coding region. There is increasing evidence, however, that 5′ and 3′ flanking sequences also contribute to tRNA promoter function. Alteration of normal 5′ flanking sequences diminishes or abolishes transcription of *Drosophila*, yeast, mouse, human and silkworm tRNA genes, showing that elements in this region provide positive transcription signals (1, 2, 3, 4, 5, 6, 7, 8, 9, 10). Moreover, flanking sequences must be responsible for the fact that members of tRNA gene families that contain identical coding sequences frequently exhibit different transcription efficiencies. In several such cases, the dependence of transcription efficiency on upstream sequences has been demonstrated directly (11, 12). Upstream sequences that act negatively have also been identified. (11, 13, 12, 14, 10). Thus, the functional significance of upstream sequences is not in doubt. What is still unclear, however, is whether a recognizable sequence motif is responsible for the function of upstream promoter elements. In particular, is a well-defined, short sequence element responsible for the positive effects of upstream promoter elements?

Unlike the highly conserved A- and B-block elements of the downstream promoter, there are no obvious sequence similarities in the 5′ flanking regions of tRNA genes from diverse species. Within certain species, however, some common sequence motifs have been recognized. In *Xenopus*, a consensus sequence, AAAGT, has been derived by comparison of the upstream regions of several 5S genes (15, 16). A similar sequence also occurs in a $tRNA^{Met}$ gene (17). Oligonucleotides corresponding to the 5′ flanks of these 5S and $tRNA^{Met}$ genes can direct the formation of a protein/DNA complex (17). However, methylation interference patterns suggest that nucleotides both within and outside of the AAAGT consensus are required for protein binding. Thus, the relationship between the AAAGT element and promoter function is not clear.

In *Drosophila*, 13 tRNA genes have been found to contain a pentanucleotide of the general sequence TNNCT in the −25 to −45 region, whereas another 10 tRNA genes lack it. There is a significant correlation between the presence of this sequence and high template activity (18). Moreover, the functional importance of specific bases within this motif (the first T and the last C,T) has been demonstrated by the effects of point mutations on transcriptional efficiency (18). As in the *Xenopus* case, however, it appears that other upstream sequences also contribute to promoter function. These may account for the transcriptional activity of certain of the *Drosophila* tRNA genes that lack the TNNCT motif.

The situation in yeast is more complicated. Failure to see a major effect of 5′ flanking sequence changes on transcription efficiency (19), together with evidence for sequence non-specific protein–DNA interactions in the 5′ flanking region of several yeast tRNA genes (20), suggests that specific upstream sequences are not essential for transcription. However, some 5′ flanking substitutions do result in decreased promoter efficiency (19, 20). Moreover, genes coding for a number of abundant yeast tRNAs ($tRNA^{Leu}$, $tRNA^{Arg}$, $tRNA^{Lys}$, $tRNA^{Ser}$) contain a conserved motif YYCAACAAATAAGT in their 5′ flanks (21). Deletion of a DNA segment containing this element from $tRNA^{Leu}$ and $tRNA^{Tyr}$ genes results in decreased transcription (22, 5).

A particularly clear example of transcriptional dependence on upstream sequences is provided by the class III genes of the

*To whom correspondence should be addressed at: Institute of Molecular Biology, University of Oregon, Eugene, OR 97403-1229, USA

silkworm, *Bombyx mori*. This paper focuses on identification of the critical sequence elements in the upstream promoter of one of these, a $tRNA_c^{Ala}$ gene. Two properties of this gene make it favorable for such an analysis. 1) Transcription of the $tRNA_c^{Ala}$ gene is very strongly dependent on the presence of natural 5′ flanking DNA. Without this sequence, transcription is undetectable (9). Thus, the system possesses the sensitivity required to detect mutant phenotypes unambiguously. 2) DNA upstream of the $tRNA_c^{Ala}$ gene contains elements that were already suspected of functional importance, based on their presence in other silkworm class III promoters. These are two AT-rich elements whose consensus sequences are TATAT and AATTTT. One or both occur at similar positions in another $tRNA_c^{Ala}$ gene (23), four $tRNA^{Gly}$ genes (10, 24), a 5S RNA gene (25), and a Bmx (=Bm1) gene (26).

An additional motivation for our work is the fact that an upstream promoter element is responsible for the unusual transcriptional properties of the *B.mori* $tRNA_{SG}^{Ala}$ gene, which is tissue-specifically expressed (27). Thus, identification of the particular sequences with promoter function for the constitutively expressed $tRNA_c^{Ala}$ gene is essential for defining the key functional distinctions between regulated and constitutive class III promoters in silkworms.

Initial studies aimed at delineating the promoter of the $tRNA_c^{Ala}$ gene had established that a region more than 11 bp upstream of the transcription start site was critical for function (9). The approximate upstream boundary of the critical element was placed at −34 by the properties of deletion mutants entering the $tRNA_c^{Ala}$ gene from the 5′ side (28). The approach we have used here was to systematically mutate the important region in 4−5 bp segments, using site-directed mutagenesis. The location of convenient restriction sites focused our effort on the region between −34 to −15.

METHODS

Mutagenesis

The parental wild type gene used in this work was a $tRNA_c^{Ala}$ gene from which the normal 5′ flanking sequences had been deleted to position −34 (28). The transcription phenotype of this gene is the same as that of a $tRNA_c^{Ala}$ gene containing 222 bp of wild type 5′ flanking DNA (28). An *EcoRI* fragment bearing the $tRNA_c^{Ala}$ gene (positions −34 to +215 with respect to the transcription initiation site) was inserted at the *EcoRI* site of a vector, pUC13M. This vector is a derivative of pUC13 in which the *HindIII* site of the polylinker has been replaced by an *MluI* site in order to allow cloning manipulations involving the *HindIII* site at −11 of the $tRNA_c^{Ala}$ gene. Small 4−5 base pair substitutions were made by oligonucleotide-directed mutagenesis in the region from −34 to −15. The four mutagenized subregions are designated A, B, C, and D in Figure 1. Oligonucleotides with degeneracies at specific positions and with appropriate restriction sites at their ends were synthesized by the Biotechnology Laboratory at the University of Oregon. The non-coding strands of these oligonucleotides are shown in Figure 1. To facilitate further mutagenesis, one family of oligonucleotides was designed to include some with an *NsiI* site (ATGCAT) near the middle of the mutagenized segment (Fig. 1, oligonucleotide C). Since introduction of this *NsiI* site did not affect transcription activity (Fig. 2a), a plasmid bearing an *NsiI* site-containing version of the $tRNA_c^{Ala}$ gene (designated pCN for mutation in region C containing *NsiI* site) was used for construction of other

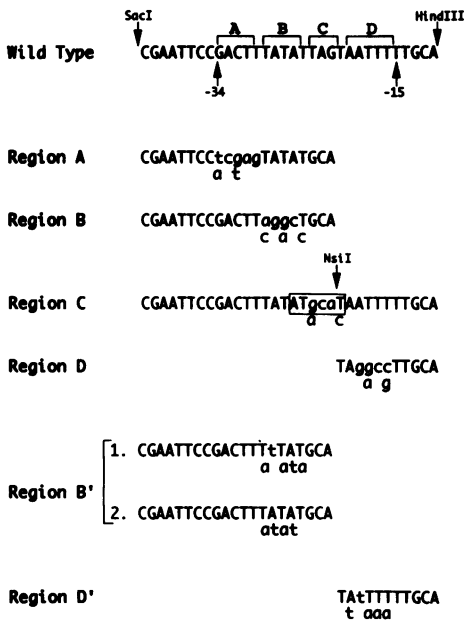


Figure 1. Oligonucleotides used for constructing mutant $tRNA_c^{Ala}$ genes. The four regions mutagenized are lettered A, B, C, and D. The non-coding strand of the wild type sequence and the oligonucleotides used to replace it are shown. Lower case letters represent bases that differ from wild type. Two letters at the same position indicate that the oligonucleotide was synthesized with a 1:1 mole ratio of two nucleotides at a single position.

mutants (Fig. 1). The oligonucleotides were inserted either between the *SacI* and *NsiI* sites or between the *NsiI* and *HindIII* sites. Mutants with sequence alterations in both AT elements were obtained by ligating oligonucleotides corresponding to regions B and D prior to insertion between the *SacI* and *HindIII* sites of pCN.

Mutant genes containing variant AT-rich elements were constructed by inserting the B′ and D′ oligonucleotides shown in Figure 1 between the appropriate restriction sites in pCN. Degeneracy at the indicated positions is expected to yield a total of 16 different sequences for each element. We focused on the subset of sequences that contains the same ratio of A to T as in the wild type elements. This set included six mutations in the AATTTT box and five in the TATAT box. Since the last T of the TATAT sequence forms part of the required overhang of the *NsiI* site, it could not be altered in this mutagenesis strategy. This position was mutated by using an alternative strategy in which mutant oligonucleotides containing a *SacI* site on one end and a blunt end at the other (Fig. 1 oligonucleotide B'1) were inserted in pCN that had been treated with T4 DNA polymerase to make the *NsiI* site blunt. This strategy yielded three mutants. We also analyzed a few mutants that did not contain the wild type A to T ratio. These were ATATA and TTTTT in the TATAT box and TTTTTT in the AATTTT box. In all cases, complementary single stranded oligonucleotides were annealed, ligated to appropriately restricted vector, and the DNA was introduced into *E.coli* (TB1) cells by electroporation.

In vitro transcription

The transcription phenotype of each mutant was determined with two different preparations of template and with two different batches of silkgland nuclear extract. Template DNA was in the

form of supercoiled plasmid that had been purified by equilibrium density sedimentation in the presence of CsCl and ethidium bromide (29). To determine the conditions necessary for sensitive detection of transcription phenotypes, each of the nuclear extracts was titrated with both template and non-specific DNA (30). These titrations established that the response to template DNA was linear up to 11 ng (6 fmols) per reaction mixture in the presence of sufficient non-specific DNA (pBR322) to bring the total to 200 ng DNA per reaction mixture. Transcription reaction mixtures contained the following in a volume of 20 ml: 5.5 ng (3 fmol) template and 200 ng total DNA, 5 μl of extract, 70 mM KCl, 4.5 mM $MgCl_2$, 28 mM Tris HCl (pH 7.5), 1mM dithiothreitol, 10% glycerol, 600 μM each of GTP, ATP and CTP, 25 μM α-[^{32}P]-UTP (10 Ci/mmol). Reactions were carried out at room temperature (~22°C) for 2h, stopped, fractionated on polyacrylamide gels, and autoradiographed as described (25).

Transcription initiation and termination sites

Oligonucleotides containing the 5′ and 3′ ends of primary transcripts were identified in T1 RNase fingerprints of *in vitro* transcription products labeled with α-[^{32}P]-UTP (31, 23). Secondary and tertiary RNA sequence analysis confirmed these identities and verified the terminal nucleotides of the primary transcripts (23).

Comparison of sequences upstream of tRNA genes

Sequences of eukaryotic tRNA genes were obtained from the GCG data base. Pseudo-genes and genes whose sequences had not been determined to −50 were eliminated. All of the others were manually aligned with respect to the first nucleotide of the mature tRNA (position +1). Using a word processing program on an IBM computer, the sequences from −1 to −50 were grouped into five files corresponding to: Insects, Fungi, Protozoa, Plants, and Vertebrates. These files were then analyzed by a spread sheet program, using a logic function to assign a value of 1 to positions containing A or T, and a value of 0 to positions containing G or C. The percent A+T for each segment of five adjacent positions from −3 to −48 was then computed. The statistical significance of the data was determined by calculating the number of standard deviations between the A+T content of each 5 bp segment and the average A+T content of genomic DNA (a Z-test). The A+T content of genomic DNA is a weighted average for the group of organisms, calculated from data in (32).

RESULTS

Critical promoter elements in the upstream region

To determine which sequence(s) within the $tRNA_c^{Ala}$ upstream promoter are critical for transcription, we systematically replaced 4−5 bp segments within this region. These segments are lettered A to D in Figures 1, 2 and 3. The replacement sequences were designed to introduce major sequence changes that typically altered the base composition of the segment, but did not disrupt the spatial relationships of adjacent sequences. As described in Methods, oligonucleotide directed mutagenesis was facilitated by the introduction of an *NsiI* site near the middle of the 5′ flanking region (region C). As shown in Figure 2a, the sequence alteration required to create this site did not affect transcription efficiency. Fingerprint analysis of transcripts synthesized from this template established that the site of transcription initiation was also unchanged (data not shown). Therefore, the $tRNA_c^{Ala}$ gene

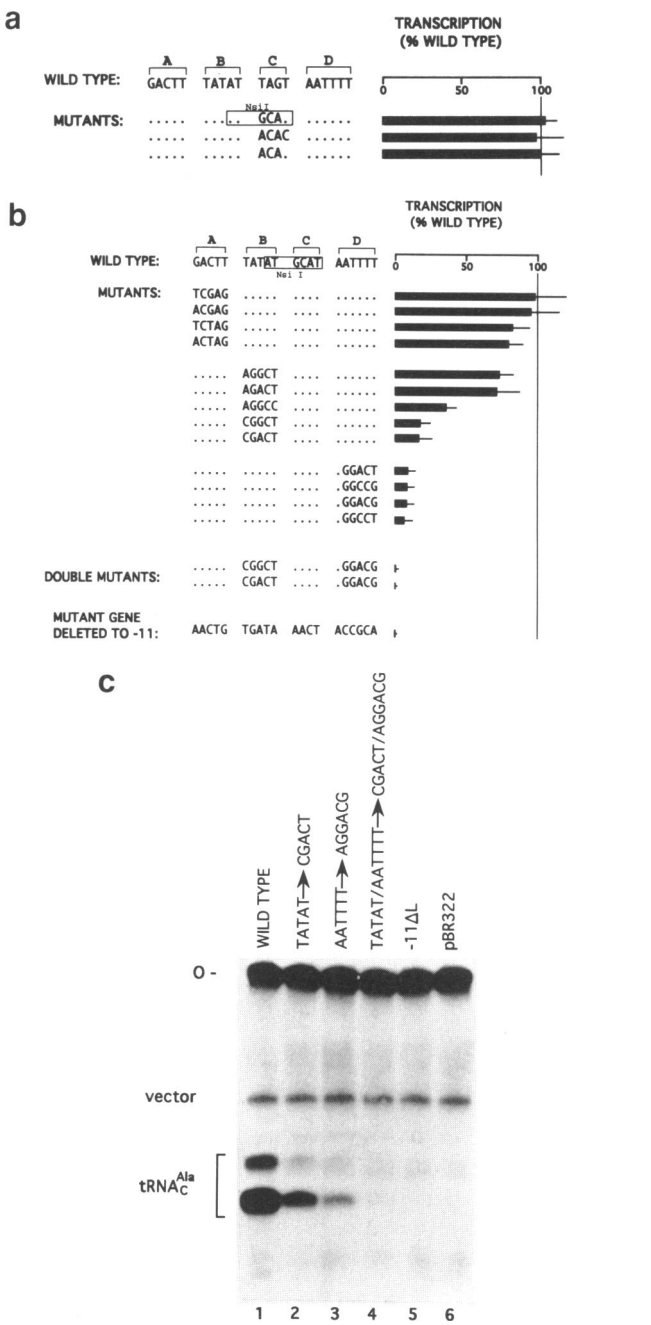


Figure 2. Transcription activity of upstream promoter mutants. The four regions mutagenized are marked A, B, C and D. The transcription activity is represented by horizontal bars. Each bar represents the mean, and each thin line the standard deviation, of four or more determinations. The vertical line at 100% represents the activity of the wild type gene. Nucleotide substitutions are shown below the wild type sequence; unchanged positions are represented by dots. **a.** Transcription activity of three different genes containing substitutions in region C. The mutation that results in the formation of an *NsiI* site, ATGCAT, is boxed. **b.** Transcription activity of genes containing substitutions in regions A, B and D. The last line shows the sequence of pBR322 that replaces silkworm DNA in a mutant lacking the entire upstream promoter (9). **c.** A representative gel showing *in vitro* transcription products from mutants in regions B and/or D. The DNA templates and their average transcription activities (% of wild type) are as follows: lane 1: wild type $tRNA_c^{Ala}$ gene (100%); lane 2: TATAT block mutant, CGACT (17 ± 9%); lane 3: AATTTT block mutant, AGGACG (8 ± 5%); lane 4: TATAT+AATTTT blocks double mutant CGACT+AGGACG (0.5 ± 1%); lane 5: a mutant deleted to −11, 5′Δ−11L (0.5% ± 1); lane 6, pBR322 (0.5% ± 1), respectively. Each value represents the mean ± SD of four or more determinations. Transcription products were run on an 8% polyacrylamide gel. 'O' denotes the origin, '$tRNA_c^{Ala}$' the position of $tRNA_c^{Ala}$ transcripts, and 'vector' denotes transcripts directed by vector DNA.

containing the *NsiI* site was used for all subsequent mutagenesis.

Figure 2b summarizes quantitatively the transcription activities of the complete set of mutants in regions A, B and D. Figure 2c shows autoradiographic data for representative mutants in regions B and D. Mutations in regions A and C (Figures 2a and 2b) had no significant effect on transcription activity, whereas mutations in regions B and D reduced it. Regions B and D correspond to the AT-rich blocks, TATAT and AATTTT, respectively, and mutation of either block, individually, weakens the promoter. Substitutions in the TATAT block (region B) reduce transcription efficiency to values between 65% and 15% of the wild type level. Mutations in the AATTTT block (region D), the element closest to the transcription start site, are even more deleterious. All of them reduce transcription to less than 10% of the wild type level. Since promoter elements might control the transcription start site, as well as the efficiency of transcription, we examined the ends of the transcripts produced by two mutant genes. We chose one mutant in each of the AT-rich elements that had a strong effect on transcription efficiency. These were, CGGCT (18.2 ± 7% of the wild type level) in the TATAT block and AGGACG (8.4 ± 5% of the wild type level) in the AATTTT block (Fig. 2b). Analysis of the 5′ and 3′ ends of these transcripts by RNase T1 fingerprinting and secondary RNA sequencing showed that transcription initiated and terminated at the wild type sites. To determine whether the combined action of the two AT-rich elements could account for all of the activity of the upstream promoter segment, we constructed mutant genes in which both AT-rich elements were altered. As shown in Figures 2b and 2c, these mutants give no detectable transcripts.

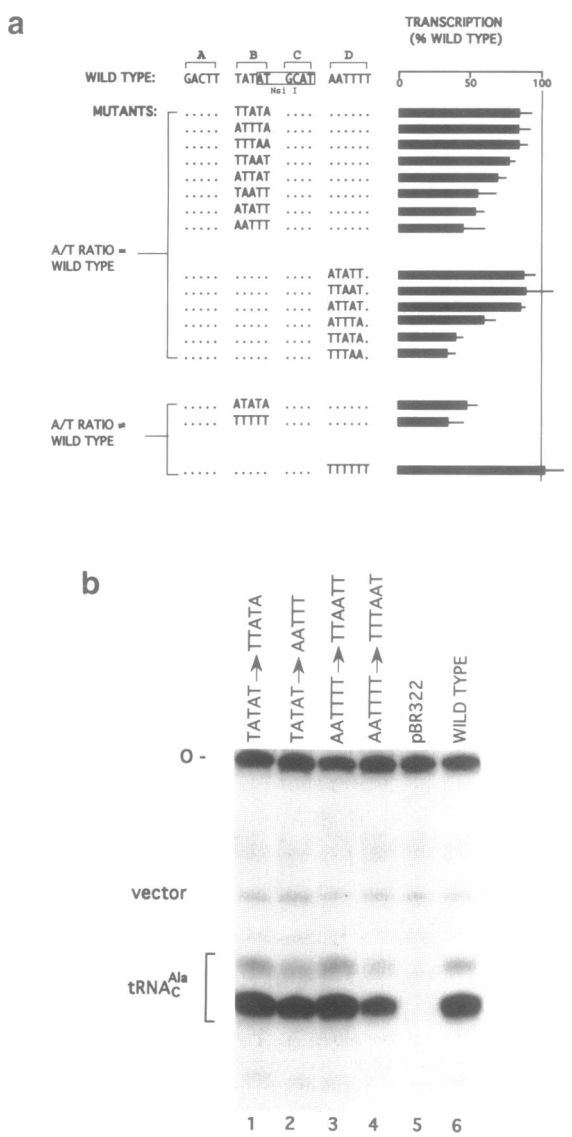


Figure 3. Transcription activity of genes with variant AT-rich elements. **a.** Transcription activities of genes that have the wild type A/T ratio (top 14 lines) or a different A/T ratio (bottom 3 lines) are shown. Symbols as in legend to Figure 2. **b.** A representative gel showing the *in vitro* transcription products from genes with variant A+T-rich elements. Mutants that exhibit the maximum and minimum activities observed for each element are shown. The DNA templates and their average transcription activities (% of wild type) are as follows: lane 1: TATAT block mutant, TTATA (84 ± 9%); lane 2: TATAT block mutant, AATTT (45 ± 15%); lane 3: AATTTT block mutant, TTAATT (88 ± 19%); lane 4: AATTTT block mutant, TTTAAT (34 ± 6%); lane 5: pBR322 (0.5% ± 1%); and lane 6: wild type $tRNA_c^{Ala}$ gene (100%). Symbols as in legend to Figure 2.

Sequence specificity *versus* base composition

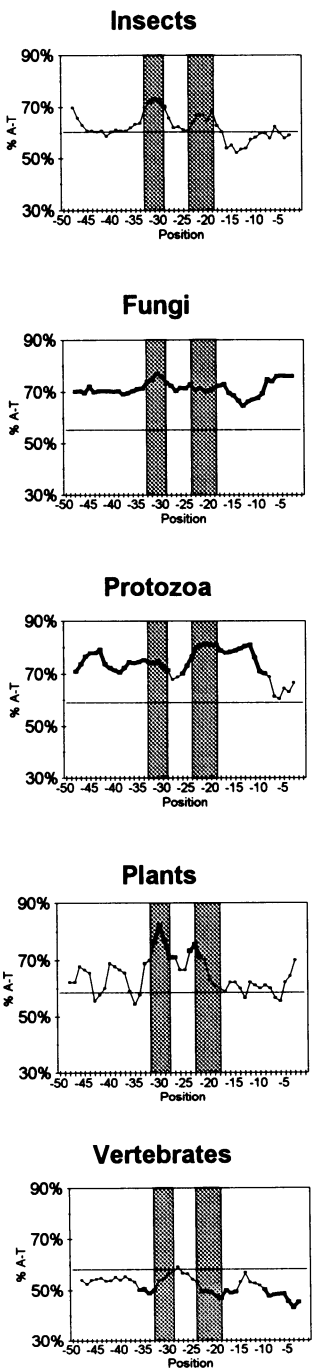
Since an obvious structural feature of the critical promoter elements is their unusual base composition, we wanted to determine whether the promoter activity of these elements requires a specific sequence, or simply a high proportion of A+T. We therefore created mutants in which the AT blocks varied in sequence, but not in A+T/G+C bias. To make this analysis both systematic and manageable in scope, we focused on the subset of mutant genes that contains the wild type ratio of A to T. A few additional mutants, which did not fit this criterion, were also analyzed. These were the complement to the TATAT block (ATATA), and runs of Ts in either block.

As shown in Fig. 3, several different AT-rich sequences in both blocks permit nearly wild-type levels of transcription. Four of the 8 TATAT mutants and 3 of the 6 AATTTT mutants are transcribed at least 75% as efficiently as are wild type templates. The strongest effect of any of the AT sequence rearrangements was a reduction to 35% of the wildtype level in one of the AATTTT mutants (TTTAAT). Thus, it is clear that a specific sequence is not essential for function of the upstream promoter elements. Certain limitations on sequence variation do exist, however. Since some replacement sequences in each of the AT elements reduce transcriptional efficiency, all AT-rich sequences cannot be functionally equivalent.

A+T-rich sequences are common in the 5′ flanks of tRNA genes

The fact that upstream AT boxes provide essential tRNA promoter function in silkworms prompted us to ask whether similar elements are associated with tRNA genes in other organisms. Although previous comparisons of sequences upstream of tRNA genes have not revealed a clear consensus [for instance, see Bertling *et al.* (33), who found only a PyPuPuPuPuPy box at variable positions], we were encouraged to look again by our finding that multiple AT-rich sequences are functionally equivalent. We reasoned that it might be AT-richness, rather than a particular sequence, that has been conserved. Therefore, we examined the upstream regions of 282 tRNA genes from 20 different species to determine whether they contain A+T-rich segments at particular locations. The base composition of 50 base pairs of 5′ flanking DNA was

systematically examined by calculating the percent A+T within 5 base pair segments centered at each position from −48 to −3. For purposes of comparison, the 20 species examined were put into five groups: insects, fungi, protozoa, plants, and vertebrates. The averaged data for each group are plotted in Figure 4. The values that are significantly different from the A+T content of genomic DNA for the group as a whole are indicated by larger dots connected by heavier lines. Figure 4 shows that AT-richness upstream of tRNA genes is common to many organisms. All groups except vertebrates have AT-rich sequences in approximately the same locations as the silkworm $tRNA_c^{Ala}$ AT boxes (centered at −30 and −20). Protozoa and fungi also have a third region of AT-richness even closer to the gene.



DISCUSSION

We have specifically mutated short segments of the $tRNA_c^{Ala}$ 5′ flanking region between −34 and −15 and have shown that mutations that alter both the sequence and the base composition of either of the two AT elements lead to reduced transcriptional activity. These mutations do not affect the choice of the transcription initiation site, however. Sequence replacements elsewhere in the upstream region have little or no effect on transcription. The decreased activity of the AT element mutants is unlikely to be due to the introduction of a particular inhibitory sequence. A variety of sequence substitutions have similar negative effects. Moreover, none of the substituted sequences is predicted to form an unusual DNA structure, such as the Z conformation that is thought to reduce transcription of a *Xenopus* $tRNA^{Met}$ gene (14). Thus, the AT elements appear to be the functionally important elements of the silkworm $tRNA_c^{Ala}$ 5′ flanking DNA. Each is necessary for full transcriptional activity, and removal of both of them eliminates the activity of the upstream promoter.

Are certain positions within these elements particularly critical for transcriptional activity? Although we do not know the effects of systematic alteration of individual base pairs, it is useful to compare multiply mutant versions of a single element that differ in transcriptional activity. Such a comparison is possible in the case of the TATAT element, and it suggests that the first and last Ts are especially important. In the three least active mutants, one or the other of these Ts is replaced by C. The fact that the first T can be changed to A with little effect on transcription suggests that the requirement may simply be for an A:T basepair, rather than a G:C basepair at these two positions. In the case of the AATTTT element, since all of the mutations reduce activity to about the same low level (~10% of wild type), the functional significance of particular positions is not apparent.

Having established the importance of the AT elements by mutations that altered both sequence and base composition, we tested whether a specific sequence, or simply AT-richness, is the important feature of the wildtype elements. At one extreme, it is possible that the AT elements are sequence-specific binding sites for some component(s) of the silkworm transcription machinery. If so, constraints on the allowable sequences are expected. At the other extreme, the AT elements could facilitate

Figure 4. AT-rich elements upstream of eucaryotic tRNA genes. Sequences of 282 tRNA genes obtained from the GCG database, from 20 different species, were grouped into the five categories shown. The species included and the number () of tRNA genes from each of them were: **Insects:** *Drosophila melanogaster* (49), *Bombyx mori* (4); **Fungi:** *Saccharomyces cerevisiae* (52), *Schizosaccharomyces pombe* (27); **Protozoa:** *Dictyostelium discoideum* (34), *Tetrahymena thermophila* (5), *Trypanosoma brucei* (7); **Plants:***Arabidopsis thaliana* (8), *Phaseolus vulgaris* (2), *Oryza sativa* (2), *Glycine max* (1), *Nicotiana rustica* (3), *Triticum aestivum* (2); **Vertebrates:** *Bos taurus* (2), *Gallus gallus* (5), *Homo sapiens* (35), *Mus musculus* (21), *Oryctolagus cuniculus* (2), *Rattus norvegicus* (16), *Xenopus laevis* (5). For each gene, the percent A+T was computed within every segment of 5 adjacent positions centered at −48 through −3, aligned with respect to the 5′ nucleotide of the mature tRNA . Each point represents the mean group base composition for the segment centered at the plotted position. Horizontal lines represent the average base composition of total genomic DNA for the groups. Segments whose base composition differs by more than 2 standard deviations from that of the group are indicated by larger dots connected by thicker lines. Vertical shaded areas correspond to the location of AT boxes in the *B.mori* $tRNA_c^{Ala}$ gene analyzed here.

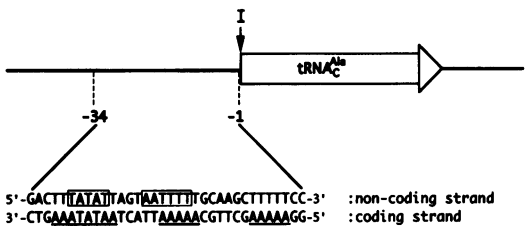


Figure 5. A-tracts upstream of the *B.mori* tRNA$_c^{Ala}$ gene. The sequence immediately upstream from a tRNA$_c^{Ala}$ gene is shown, numbered from the site of transcription initiation (I). A-tracts in the coding strand are underlined. The transcriptionally important AT elements are boxed in the non-coding strand.

localized melting *via* the thermodynamic instability of A:T basepairs, and, hence, promote the formation of a transcription bubble. In this case, a variety of AT-rich sequences should be functionally equivalent.

The observed result is that several different AT-rich sequences provide nearly wildtype function for both elements. Thus, a particular sequence is not essential for promoter function. At first glance, this result seems to argue against a role for the AT elements as protein binding sites. On the other hand, the fact that the various AT-rich sequences are not equally active argues against a purely thermodynamic role. A possible resolution of this paradox is suggested by the properties of proteins that are known to bind AT-rich DNA. Proteins such as IHF (34), TBP (35, 36) and the HMG proteins [reviewed by Churchill and Travers (37)] bind AT-rich DNA by contacting the minor groove. Unlike the major groove, the minor groove offers little in the way of distinctive hydrogen bonding patterns to discriminate between AT and TA base pairs (38). These proteins, therefore, are predicted to recognize a variety of AT-rich sequences. In fact, binding of yeast and human TBP to non-consensus sequences has been documented (39, 40). Moreover, the importance of AT-richness in sequences that are functional binding sites is clear in both of these systems. Wobbe and Struhl (39) showed that although not all A to T or T to A substitutions in the yeast-TATA element function equivalently, replacements by G or C residues are most deleterious. Similarly, mutations in the SV40 major late promoter showed that both TBP binding and transcription activity are enhanced when C is replaced by T, whereas both activities are reduced when T residues are replaced by C or G (40).

The latitude of allowed sequences within the silkworm AT elements is consistent with the idea that these elements interact with a minor groove binding protein. Are there any candidates for such a protein? Unpublished results point to a component (or components) of the TFIIIB fraction of the silkworm class III transcription machinery. Competition for silkworm TFIIIB by wildtype or mutant genes has established that TFIIIB requires a normal 5′ flanking sequence for interaction with the tRNA$_c^{Ala}$ gene (H.Sullivan, L.Young and K.Sprague, unpublished). In contrast, TFIIIC does not depend on this sequence for interaction (41). An attractive hypothesis is that TATA-binding protein, TBP, mediates the interaction of TFIIIB with the tRNA$_c^{Ala}$ upstream promoter. TBP is required for transcription of tRNA or tRNA-like genes in other systems (42, 43, 44), and is present in typical TFIIIB fractions in these systems (45, 46, 47).

Although neither of the silkworm AT elements possesses the consensus TATA sequence derived from silkworm class II genes, TATAAAA (48, 49), both resemble it. As shown below, the non-coding strand of the TATAT element is similar to consensus, and part of the coding strand of this element contributes to a sequence that differs from consensus by only one nucleotide.

non-coding: 5′ GACTT**TATAT**TAGTAATTTTTGCA 3′
coding: 3′ CTG**AAATAT**AATCATTAAAAACGT 5′

The idea that TBP acts through upstream tRNA promoter elements is consistent with the presence of AT elements upstream of tRNA genes from a wide range of organisms. Although not yet directly demonstrated, it seems likely that TBP is a universal component of class III transcription machinery. We hypothesize that in many organisms, TBP interacts directly with tRNA promoter elements through AT-rich sequences. Vertebrates constitute an informative exception to this generalization. The notable absence of AT-richness upstream of mammalian tRNA genes correlates with the fact that adaptor proteins (TAFs) are required for TBP function on human class III genes (45, 46).

How might the AT-rich elements facilitate specific interaction with TBP, or a TBP-like polypeptide? As indicated above, if the chief contacts are in the DNA minor groove, complementary hydrogen bonding between the protein and DNA may not make a major contribution to site-specific binding. On the other hand, the AT-rich elements may promote an alteration in DNA structure that facilitates protein binding. In this light, it is interesting that the known AT-rich DNA binding proteins create a bend in their target sites (50, 51). Therefore, the capacity for intrinsic or induced bends may be the essential feature of such sites, and differences in the ability to bend may account for differences in activity. In support of this idea, recent crystallographic data show that the DNA in TBP-TATA complexes is sharply (100°) bent (52, 53). Moreover, unpublished results indicate that the degree of bending caused by TBP binding to variant class II TATA boxes generally correlates with the ability of these sites to direct basal transcription (D.B.Starr, B.R.Hoopes, and D.K.Hawley, pers. commun.). The relevance of these observations to class III transcription is suggested by the fact that a yeast fraction that includes TBP, the TFIIIB fraction, causes a bend near the transcription start site of a tRNA^{Gln} gene (54) and a 5S rRNA gene (55).

With these observations in mind, we asked whether the sequences upstream of a wildtype tRNA$_c^{Ala}$ gene are predicted to be flexible or bendable, and whether mutations that reduce transcription would be expected to interfere with bending. The full upstream sequence of the tRNA$_c^{Ala}$ gene, from −1 to −34, is shown in Figure 5. As indicated in the figure, the coding strand of this segment contains two uninterrupted A-tracts and one interrupted A-tract whose spacing (10−11 nucleotides between the 5′ terminal A residues) and length (5 nucleotides) should favor bending (56). These A tracts might act by providing an intrinsic bend, or by reducing the energy required to induce a bend upon protein binding. The A-tract that overlaps the AATTTT element in the non-coding strand, 5′CAAAAAT3′, is likely to be especially effective at promoting bending because both its length and its flanking nucleotides, 5′C and 3′T, are optimal (56). The decreased transcription activity associated with substitution of A for T at position 5 in the AATTTT element is potentially explained by a decrease in bending due to disruption of this A tract.

In conclusion, our results show the importance of AT elements in tRNA promoter function in silkworms, and suggest that such elements might promote class III transcription in other organisms, as well. Current work is aimed at determining whether silkworm

TBP interacts with the $tRNA_c^{Ala}$ AT elements, and, if so, how that interaction facilitates transcription.

ACKNOWLEDGMENTS

We are grateful to Debra McMillen and the University of Oregon Biotechnology Laboratory for synthesis of oligonucleotides, Connie N.White for construction of pUC13M, Ching Ouyang and Poh Kheng Loi for help with mutant isolation, and all the members of the laboratory for helpful discussions and critical reading of the manuscript. This work was supported by National Institutes of Health research grant GM25388 to K.U.S.

REFERENCES

1. Schaack, J., Sharp, S., Dingermann, T., Burke, D.J., Cooley, L. and Söll, D. (1984) *J. Biol. Chem.*, **259**,1461–1467.
2. Sajjadi, F.G., Miller, R.C., Jr. and Spiegelman, G.B. (1987) *Molec. Gen. Genet.*, **206**, 279–284.
3. Lofquist, A.K. and Sharp, S.J. (1986) *J. Biol. Chem.*, **261**, 14600–14606.
4. Raymond, G.J. and Johnson, J.D. (1983) *Nucl. Acids Res.*, **11**, 5969–5988.
5. Shaw, K.J. and Olson, M.V. (1984) *Mol. Cell Biol.*, **4**, 657–665.
6. Morry, M.J. and Harding, J.D. (1986) *Mol. Cell. Biol.*, **6**, 105–115.
7. Rooney, R.J. and Harding, J.D. (1988) *Nucl. Acids Res.*, **16**, 2509–2521.
8. Arnold, G.J. and Gross, H.J. (1987) *Gene*, **51**, 237–246.
9. Sprague, K.U., Larson, D. and Morton, D. (1980) *Cell*, **22**,171–178.
10. Fournier, A., Guerin, M.-A., Corlet, J. and Clarkson, S.G. (1984) *The EMBO J.*, **3**,1547–1552.
11. DeFranco, D., Schmidt, O. and Söll, D. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3365–3368.
12. Dingermann, T., Burke, D., Sharp, S., Schaack, J. and Söll, D. (1982) *J. Biol. Chem.*, **257**, 14738–14744.
13. De Franco, D., Sharp, S. and Söll, D. (1981) *J. Biol. Chem.*, **256**, 12424–12429.
14. Hipskind, R.A. and Clarkson, S.G. (1983) *Cell*, **34**, 881–890.
15. Korn, L.J. and Brown, D.D. (1978) *Cell*, **15**,1145–1156.
16. Nietfeld, W., Digweed, M., Mentzel, H., Meyerhof, W., Köster, M., Knöchel, W., Erdmann, V.A. and Pieler, T. (1988) *Nucl. Acids Res.*, **16**, 8803–8814.
17. Oei, S.-L. and Pieler, T. (1990) *J. Biol. Chem.*, **265**, 7485–7491.
18. Sajjadi, F.G. and Spiegelman, G.B. (1987) *Gene*, **60**,13–19.
19. Koski, R.A., Allison, D.S., Worthington, M. and Hall, B.D. (1982) *Nucl. Acids Res.*, **10**, 8127–8143.
20. Kassavetis, G.A., Riggs, D.L., Negri, R., Nguyen, L.H. and Geiduschek, E.P. (1989) *Mol. Cell. Biol.*, **9**, 2551–2566.
21. Raymond, K.C., Raymond, G.J. and Johnson, J.D. (1985) *EMBO J.*, **4**, 2649–2656.
22. Johnson, J.D. and Raymond, G.J. (1984) *J. Biol. Chem.*, **259**, 5990–5994.
23. Morton, D.G. and Sprague, K.U. (1982) *Mol. Cell. Biol.*, **2**,1524–1531.
24. Taneja, R., Gopalkrishnan, R. and Gopinathan, K.P. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1070–1074.
25. Morton, D.G. and Sprague, K.U. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 5519–5522.
26. Wilson, E.T., Condliffe, D.P. and Sprague, K.U. (1988) *Mol. Cell. Biol.*, **8**, 624–631.
27. Young, L.S., Takahashi, N. and Sprague. K.U. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 374–378.
28. Larson, D., Bradford-Wilcox, J., Young, L.S. and Sprague, K.U. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3416–3420.
29. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1989) Current Protocols in Molecular Biology. John Wiley & Sons, New York.
30. Wilson, E.T., Larson, D., Young, Y.S. and Sprague, K.U. (1985) *J. Mol. Biol.*, **183**,153–163.
31. Platt, T. and Yanofsky, C. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 2399–2403.
32. Fasman, G.D. (ed.) (1975) Handbook of Biochemistry and Molecular Biology: Nucleic Acids, Vol II, CRC Press, Ohio.
33. Bertling, W., Dingermann, T. and Kaiserwerth, M. (1987) *Intl. J. Biol. Macromol.*, **9**, 63–70.
34. Yang, C.-C. and Nash, H.A. (1989) *Cell*, **57**, 869–880.
35. Starr, B.D. and Hawley, D.K. (1991) *Cell*, **67**,1231–1240.
36. Lee, D.K., Horikoshi, M. and Roeder, R.G. (1991) *Cell*, **67**, 1241–1250.
37. Churchill, M.E.A. and Travers, A.A. (1991) *Trends Biochem. Sci.*,**16**, 92–97.
38. Seeman, N.C., Rosenberg, J.M. and Rich, A. (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 804–808.
39. Wobbe, R.C. and Struhl, K. (1990) *Mol. Cell. Biol.*, **10**, 3859–3867.
40. Wiley, S.R., Kraus, R.J. and Mertz, J.E. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 5814–5818.
41. Young, L.S., Rivier, D.H. and Sprague, K.U. (1991) *Mol. Cell Biol.*, **11**, 1382–1392.
42. Schultz, M.C., Reeder, R.H. and Hahn, S. (1992) *Cell*, **69**, 697–702.
43. White, R.J., Jackson, S.P. and Rigby, P.W.J. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1949–1953.
44. Cormack, B.P. and Struhl, K. (1992) *Cell*, **69**, 685–696.
45. White. R.J. and Jackson, S.P. (1992) *Cell*, **71**,1–20.
46. Lobo S.M., Tanaka, M., Sullivan, M.L. and Hernandez, N. (1992) *Cell*, **71**, 1029–1040.
47. Huet, J. and Sentenac, A. (1992) *Nucl. Acids Res.*, **20**, 6451–6454.
48. Tamura, T., Inoue, H. and Suzuki, Y. (1987) *Mol. Gen. Genet.*, **207**,189–195.
49. Okamoto, H., Ishikawa, E. and Suzuki, Y. (1982) *J. Biol. Chem.*, **257**, 15192–15199.
50. White, S.W., Appelt, K., Wilson, K.S. and Tanaka, I. (1989) *Prot. Struct. Func. Genet.*, **5**, 281–288.
51. Nash, H.A. (1990) *Trends Biochem. Sci.*, **15**, 222–227.
52. Kim, Y., Geiger, J.H., Hahn, S. and Sigler, P.B. (1993) *Nature* **365**, 512–520.
53. Kim., J.L., Nikolov, D.B. and Burley, S.K. (1993) *Nature* **365**, 520–527.
54. Léveillard, T., Kassavetis, G.A. and Geiduschek, E.P. (1991) *J. Biol. Chem.*, **266**, 5162–5168.
55. Braun, B.R., Kassavetis, G.A. and Geiduschek, E.P. (1992) *J. Biol. Chem.*, **267**, 22562–22569.
56. Koo, H.-S., Wu, H.-M. and Crothers, D. (1986) *Nature*, **320**, 501–506.