Nucleic Acids Research, Vol. 18, No. 6

Saturation mutagenesis of the *Drosophila* tRNA^{Arg} gene B-Box intragenic promoter element: requirements for transcription activation and stable complex formation

Bruno A.Gaëta, Stephen J.Sharp[1]* and Thomas S.Stewart
School of Biochemistry, University of New South Wales, PO Box 1, Kensington, 2033 Australia and [1]Department of Microbiology and Molecular Genetics, California College of Medicine, University of California, Irvine, CA 92717, USA

Received November 17, 1989; Revised and Accepted February 2, 1990

ABSTRACT

Transcription of eukaryotic tRNA genes is dependent on the A- and B-Box internal control regions (ICRs) and the upstream transcription modulatory region. The B-Box ICR spans nucleotides 52 to 62 and directs the primary binding of transcription factor C as the first step in the formation of a transcription complex. The conservation of the sequence of the B-Box in all tRNA species reflects its importance in both the expression of the gene and the processing, structure and function of the gene product. In order to identify the nucleotides essential to the promoter function of the B-Box ICR, site-directed mutagenesis was used to generate all the possible single point mutations at positions 52 to 58, 61 and 62 of a *Drosophila melanogaster* tRNA^{Arg} gene. The effect of these mutations on gene transcription was evaluated using *in vitro* transcription and template exclusion competition assays. Optimal activity was displayed by the wild type tDNA^{Arg} B-Box sequence but several other sequences supported *in vitro* transcription at wild type levels. The majority of mutants, however, showed lower efficiency in the *in vitro* transcription assay. Of the single point mutations, those at positions 53, 55, and 56 had a critical effect on gene function in *Drosophila* and HeLa transcription extracts and transcription factor interaction most likely requires base contacts at these positions. Since the effect of several of the point mutations cannot be explained in terms of possible major or minor groove contributions the possibility is raised that local DNA geometry also is an important determinant in specifying B-Box function.

INTRODUCTION

Transcription of eukaryotic tRNA genes (tDNA) by RNA polymerase III is controlled at the DNA level by extragenic and intragenic sequences (reviewed in 1 and 2). The primary sequences that direct formation of transcription complexes are the A- and B-Box internal control regions (ICRs) which by deletion-substitution, linker-scanning, and point mutation analyses have been delimited to coordinates 8 to 19 for the A-Box, and coordinates 52 to 62 for the B-Box (1, 2). Since the initial event in establishing transcription competency of a tRNA gene involves binding of the transcription factor TFIIIC to the B-Box, the mechanism for this interaction has been a major concern in understanding transcription activation by RNA polymerase III (3–8).

TFIIIC of yeast appears to be a multimeric protein that contains two DNA binding domains, the B domain, which binds to the B-Box ICR, and the A domain, which binds to the A-Box ICR (9–12). Human TFIIIC appears to be comprised of at least two separate components, TFIIIC1 and TFIIIC2 (13, 14) and, binding of TFIIIC1 to the A-Box is dependent on the prior binding of TFIIIC2 to the B-Box (15).

The B-Box encodes the T-loop region of the familiar 'cloverleaf' secondary structure of tRNA. The sequence of the T-loop is essential for the 'L' tertiary structure of tRNA and as such is conserved in all cytoplasmic tRNA species. Because of this conservation and its importance at the RNA level, a consensus sequence for the B-Box, derived by comparing the sequence of all tRNA species, may not necessarily reflect nucleotide involvement in promoter function. In line with the notion of the derived consensus sequence, *E. coli* tRNA genes that satisfy the requirements of the consensus sequence (A-Box requirements were also satisfied) are suitable templates for transcription by RNA polymerase III (16). Also, the drastic reduction in transcription efficiency of the yeast tRNA^{Tyr} gene caused by the point mutation G_{56}, showed the importance of the conserved C_{56} in B-Box function (17). In a direct binding assay, TFIIIC failed to bind to the G_{56}-containing template (18). The point mutations A_{54}, and C_{57}, in the B-Box of the tRNA^{Tyr} gene also depart from the B-Box consensus sequence and both of these mutations decreased the binding of TFIIIC (17, 18). Changing the G_{62} of the tRNA^{Tyr} gene to the consensus C_{62} increased the binding of TFIIIC to the B-Box (18), although this mutation did not affect transcription efficiency (17). While analysis of the transcription control regions of the genes for *Xenopus* tRNA$_i^{Met}$ (19), *C. elegans* tRNA^{Pro} (20), and yeast tRNA^{Tyr} (17) and tRNA$_3^{Leu}$

* To whom correspondence should be addressed

(21), demonstrate the importance of discrete nucleotide positions in B-Box function, the full extent of the determinants that specify B-Box function are not known.

A full understanding of the mechanism of tDNA transcription activation will require both an understanding of the factors and of the specific features of the recognition element that allow a factor to discriminate this sequence from all others. One approach to the study of DNA-protein interactions is to analyze the substrate DNA to which the protein binds and determine the possible contributions of each of the elements of the sequence that specify binding (22, 23). Detailed studies of prokaryotic transcription regulators have shown the importance of complementary hydrogen-bonding of amino acids in the protein to functional groups in the major or minor grooves of the DNA double helix, in specifying DNA-protein interactions (see 24–27, and reviewed in 28).

In this study, site-directed mutagenesis has been used to generate all the possible single point mutations (G, A, T, C) at positions 52 to 58, 61 and 62, representing the B-Box of a *Drosophila melanogaster* $tRNA^{Arg}$ gene (29). Since mutations at nucleotide positions 59 and 60 in various tRNA genes have had a neutral effect on their transcription (1, 2), these positions were not examined in the present study. Transcriptional analysis of the mutant tDNAs defines the specific bases important in determining B-Box function. These results support the consensus sequence hypothesis of Hall and colleagues (6) by showing that optimal function of the B-Box ICR is achieved by the presence of the tDNA T-region consensus sequence which was derived originally by a sequence comparison of all eukaryotic tRNA sequences (30).

EXPERIMENTAL PROCEDURES

DNA and Oligonucleotides

Oligonucleotides were synthesized by OCS Laboratories (Denton, Texas) and using an Applied Biosystems 380B DNA synthesizer. Each was synthesized as a mixture of three oligonucleotides, differing at the mismatch base. The $tRNA^{Arg}$ gene insert 3.106 (31) was grown in the form of single and double stranded DNA for site-directed mutagenesis, sequencing and transcription. Single stranded pArg3.106 in M13 mp9 (*Eco*RI/*Hind*III insert) was purified using a linear alkaline sucrose gradient (2.5–20% sucrose, 30 mM NaOH, 2 mM EDTA) by centrifugation at 30,000 rpm for 2 h in a Beckman SW41 rotor. Double-stranded DNA for transcription assays was cloned into plasmid pUC8 and prepared using two rounds of CsCl gradient centrifugation. DNA concentrations were estimated by absorbance at 260 nm (1.0 A_{260} is given by 50 μg/ml of double stranded DNA at A_{260}/A_{280} of 1.8). The concentrations of DNA for transcription assays were confirmed by ethidium bromide staining after agarose gel electrophoresis.

The following oligonucleotides (listed 5′ to 3′), which anneal to the non-coding strand sequence were used to generate point mutations at the positions noted (underlined or degenerate):

52, CAGGAGTCGAAC(G/A/T)TGGAATC;
53, GAGTCGAA(G/A/T)CTGGA;
54, GAGTCGA(G/T/C)CCTGGA;
55, CCAGGAGTCG(T/C)ACCTGGA, GGAGTCGGACCTGG;
56, GCCAGGAGTC(A/C/T)AACCTGGAA;
57, GCCAGGAGT(G/A)GAACCTGGAA, CCAGGAGTTGAACC;
58, GCCAGGAGCCGAACCTGG, CCAGGAG(G/A)CGAACC;
61, CCTGCCAG(A/C/T)AGTCGAACC;
62, TCCTGCGA(A/C/T)GAGTCGAAC.

In addition, complementary oligonucleotides, 5′ GGCTCTCCTCCT 3′ and 5′ GGAGGAGAGCCT 3′, were inserted as a duplex into *Zmo*I-digested pArg3.106 (*Zmo*I was kindly provided by Dr. A. G. Mackinlay) —pArg contains a *Zmo*I site on each side of the T-stem— to construct a B-Box containing C_{54}, T_{57}, and C_{58}. This sequence also inserted in the opposite orientation to form the construct containing $A_{54}G_{55}G_{56}A_{57}G_{58}A_{59}G_{60}$, called 'reverse construct'.

Site-Directed Mutagenesis

Point mutations in the $tRNA^{Arg}$ gene were constructed using the gapped-duplex method using the large *Eco*RI/*Hind*III fragment of M13 mp9rev (32). Minor modifications made to the method included the use of the mismatch repair deficient strain *E. coli* BHM 71-18 mutL (33); the amplification step was replaced by plating aliquots of the competent E. coli BHM 71-18 mutL cell suspension after transfection directly onto a lawn of *E. coli* MK30/3 cells. Plaque lifts were screened using 5′-^{32}P-labelled oligonucleotide mix that was used in the mutagenesis and filters were washed in 6×SSC at 2°C below the T_m given by the Wallace rule (34). Plaques were selected and mutants characterized by the dideoxy method of DNA sequencing using the M13 universal primer and modified T7 DNA polymerase (35).

In vitro Transcription and Template Exclusion Competition Assays

Transcription reactions contained 200 ng template DNA (wild type or mutant pArg) and 800 ng carrier pUC8, 15 μl *Drosophila* Schneider S2 or HeLa cell extract, 30 mM HEPES-KOH, pH 8.0, 70 mM KCl, 5 mM $MgCl_2$, 500 μM each of ATP, CTP, and GTP, 100 μM [α-^{32}P]UTP (0.1–0.4 Ci/mmol), 8 mM creatine phosphate and 4 units of creatine phosphokinase in a 60 μl reaction volume (36). Reactions were incubated at 24°C and stopped after 90 min by the addition of 50 μl of a solution containing 1 mg/ml proteinase K, 0.1% (w/v) sodium dodecyl sulfate in 30 mM HEPES-KOH, pH 8.0 (preincubated at 37°C, 20 min), and incubated at 37°C for 1 h. After phenol and chloroform extraction, transcription products were collected by ethanol precipitation, resuspended in a solution containing 80% (v/v) formamide, 0.05% (w/v) xylene cyanol FF, 0.05% (w/v) bromphenol blue, 1 mM EDTA, and electrophoretically resolved using an 8% (w/v) polyacrylamide gel (20:1 acrylamide:N,N′-methylenebisacrylamide) containing 8.3 M urea, 89 mM Trizma base, 89 mM boric acid, and 2 mM EDTA (36). Transcription was quantitated by liquid scintillation counting of excised radioactive bands dissolved overnight at 45°C in 0.2 ml of $HClO_4$ (60%v/v) and 0.4 ml of H_2O_2 (30% v/v).

Template exclusion competition assays were performed as described (37). DNA I (wild type or mutant pArg) was preincubated in transcription reactions for 15 min before the addition of DNA II, pArg26×36 or pArg 3.72Term maxigenes (38, 31). Maxigene was added and transcription was allowed to proceed for an additional 90 min. In these assays the amount of maxigene transcription was quantitated.

RESULTS

Construction of Point Mutations

To demonstrate the efficacy of constructing point mutations in the *Drosophila* $tRNA^{Arg}$ gene using relatively short synthetic oligonucleotides, we had previously constructed point mutations in the D-loop (A-Box) and anticodon encoding regions using the

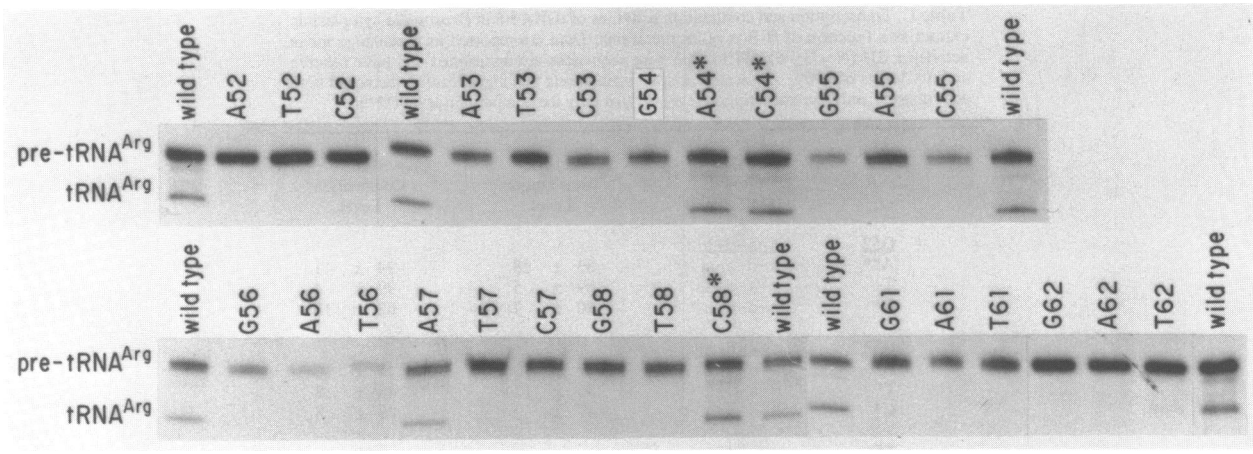

Figure 1. Effect of B-Box point mutations on in vitro transcription of a *Drosophila* tRNA^{Arg} gene in a *Drosophila* cell cytoplasmic extract. Shown is an autoradiograph of mutant transcripts analyzed on a denaturing polyacrylamide gel. Lanes marked with an asterisk indicate that the wild-type tRNA processing pattern was maintained for these mutants.

two primer method of site-directed mutagenesis (37). The use of oligonucleotides degenerate for three nucleotides at one position, in the two-primer method generally did not yield all three required mutant sequences. The gapped-duplex method was found to yield a greater percentage of mutant sequences and more importantly, this method allowed the use of mixed oligonucleotides, degenerate at one position, to generate precisely defined point mutations. We were able to use synthetic oligonucleotides of 14 to 20 nucleotides in length in the gapped-duplex method to generate all three point mutations at each position of the tRNA^{Arg} gene B-Box ICR.

B-Box Point Mutations Affect Precursor tRNA Processing

To evaluate the effects of the point mutations on transcription function, a *Drosophila* Schneider cell and HeLa cell cytoplasmic extract were used as the source of RNA polymerase III and transcription factors. Specific processing of precursor tRNA in the *Drosophila* transcription reactions was inhibited by all except three of the introduced mutations (processing activity was typically low or asbent for the HeLa extract), and thus transcription of mutant templates resulted in accumulation of a precursor-like transcript of ~90 nucleotides in length.

Processing is the production of the mature-sized tRNA by specific nucleases that act on the 5′ leader and 3′ trailer of the precursor (primary) transcription product. Most point mutations in the B-Box (T-loop encoding region) effected a decrease in the relative level of processing in the *Drosophila* extract (Fig. 1). The T-loop and -stem region of tRNA have a critical role in the tertiary folding of tRNA (39). Seemingly, most of the point mutations disrupt this structure and the resultant mutant transcripts are not suitable substrates for the processing apparatus (Fig. 1). For example, point mutations in positions 52, 53, 61, and 62 would potentially disrupt T-stem formation. Similarly, positions 55 and 56 in tRNA are involved in tertiary base interactions respectively with positions 18 and 19 in the D-loop (39). Since all tRNA species contain a purine at position 57, the observed processing of precursor-tRNA^{Arg} containing A_{57} rather than the wild type G_{57} was expected (Fig. 1). This result, combined with the template activity results (see next section), indicates that the role of the conserved purine at position 57 in tRNA is at the RNA level rather than at the transcriptional level.

Transcript RNAs containing G_{54}, G_{58} or U58 were not processed (Fig. 1). However, for transcripts containing C_{54} or C_{58} processing was qualitatively the same as for wild type (Fig. 1). In tRNA, tertiary base pairing occurs between 1-methyladenine at position 58 and 5-methyluracil at position 54 with the modifying constituents of these bases not being directly involved in this hydrogen bond interaction (39). Presumably, in the point mutants the O6 of C_{54} and N7 of A_{54} are able to mimic the wild type O6 of T_{54} to allow formation of a correct tRNA tertiary structure and participation in the processing reaction (see Fig. 1). All tRNA species contain 1-methyladenine at position 58 and T_{54} is encoded in all tRNA genes except those encoding initiator tRNA^{Met} and *Bombyx* tRNA^{Ala}, which contain A_{54} (40). The capability of the C_{58}-containing transcript to be processed (Fig. 1) can be explained by the potential of C_{58} to provide the same set of interactions as the wild-type A_{58}. The presence of C_{54} would potentially compromise tRNA structure since a second hydrogen-bond interaction between the N7 of A_{58} and the N2 hydrogen of T_{54} or the amino-C6 hydrogen of A_{54} would be absent in the point mutation C_{54}. However, this result was not observed in the present assays. From these results, the absence of C_{54} or C_{58} from wild-type tRNA species is not explained in terms of template activity (see next section) or precursor tRNA processing.

The Effect of B-Box Point Mutations on tDNA Template Activity

The relative activity of each of the mutants was determined by excising the appropriate RNA band from the gel and comparing the amount of radioactivity resulting from the mutant gene to that obtained from a reference sample of pArg3.106 (wild-type) included in the same set of reactions. Degradation of RNA is an inherent problem in this type of analysis (documented for example in 37) so, where possible, the extent and rates of transcript degradation were assessed by incubating the isolated transcript RNAs in transcription reactions and determining the amount of degradation. Degradation was enhanced in several instances by the point mutation, presumably as a result of disrupting the secondary and/or tertiary structure of the tRNA. Most notably, mutations at position 58 increased susceptibility to nuclease degradation. The extent of transcript degradation was not constant in separate experiments and this is reflected in the standard deviation of the average values (see Tables 1 and 2).

Table 1. Transcription and competition activities of $tDNA^{Arg}$ in *Drosophila* cytoplasmic extract as a function of B-Box point mutations. Data are reported as percentage mean activity ± SD (N = 3 - 5). Wild type base sequences are underlined and have relative activity levels of 100%. Indicated are point mutations that significantly decreased both activities (*), only the transcription level (**), or only the competition level (***).

tDNA	Major Groove (see Table 3)	Relative Transcription Level	Relative Competition Level
G52	a-a-d-h		
A**	a-d-a-m	83 ± 18	94 ± 11
T	m-a-d-a	89 ± 5	99 ± 4
C*	h-d-a-a	80 ± 2	82 ± 13
G53	a-a-d-h		
A*	a-d-a-m	47 ± 10	44 ± 8
T*	m-a-d-a	56 ± 12	48 ± 8
C*	h-d-a-a	47 ± 9	28 ± 8
T54	m-a-d-a		
G*	a-a-d-h	56 ± 9	27 ± 4
A	a-d-a-m	105 ± 13	98 ± 5
C	h-d-a-a	104 ± 25	111 ± 3
T55	m-a-d-a		
G*	a-a-d-h	30 ± 8	3 ± 12
A*	a-d-a-m	66 ± 4	22 ± 3
C*	h-d-a-a	32 ± 6	0 ± 3
C56	h-d-a-a		
G*	a-a-d-h	22 ± 8	19 ± 1
A*	a-d-a-m	7 ± 4	2 ± 5
T*	m-a-d-a	16 ± 4	0 ± 6
G57	a-a-d-h		
A	a-d-a-m	90 ± 18	110 ± 6
T	m-a-d-a	99 ± 26	94 ± 8
C***	h-d-a-a	100 ± 27	80 ± 12
A58	a-d-a-m		
G	a-a-d-h	90 ± 30	93 ± 4
T	m-a-d-a	90 ± 37	98 ± 6
C	h-d-a-a	117 ± 21	105 ± 5
C61	h-d-a-a		
G*	a-a-d-h	70 ± 9	84 ± 7
A**	a-d-a-m	52 ± 8	92 ± 6
T*	m-a-d-a	80 ± 6	80 ± 10
C62	h-d-a-a		
G*	a-a-d-h	81 ± 13	80 ± 11
A*	a-d-a-m	81 ± 4	73 ± 4
T**	m-a-d-a	65 ± 26	93 ± 1
C54 T57 C58		48 ± 7	30 ± 10
"Reverse Construct"		10 ± 3	14 ± 5

The effects of the various mutations on the transcriptional activity of pArg in *Drosophila* (Fig. 1; Table 1) and HeLa (Table 2) were quantitated. Using the criterion that a difference from pArg3.106 of $\geq 15\%$ is significant (after reference 18), 18 of the 27 point mutations caused decreased transcription of pArg in the *Drosophila* extract (Table 1) whereas in HeLa extract, 17 of the 27 significantly decreased the relative transcription level (Table 2); under the present assay conditions, point mutations more seriously affected transcription in the HeLa system compared to the *Drosophila* system (summarized in Figure 2). In both systems however, any base-pair change at positions 53, 55, and 56, drastically decreased transcription efficiency (Fig. 2; Tables 1 and 2). Point mutations at position 52 decreased transcription in *Drosophila* but not in HeLa extract. The mutation G_{58} led to a significant decrease in the template activity in HeLa extract but not in *Drosophila* (Tables 1 and 2). Also, while G_{62} and A_{62} led to significant decreases in the level of transcription in both systems, T_{62} only decreased the transcription level in *Drosophila* (Tables 1 and 2).

The Ability of tDNA to Stably Sequester Transcription Components

In the template exclusion competition assay, a B-Box mutant tDNA was preincubated in transcription reactions for 15 min prior to the addition of a fixed concentration of a reference template (pArg26×36 or pArg 3.72Term) and the reaction continued for an additional 90 min. The preincubation period allows time for the test gene to stably sequester transcription components. The test gene sequesters the transcription components and thereby excludes their availability for maxigene, limiting the extent of maxigene transcription (21). Quantitation of the level of maxigene transcription measures the ability of tDNAs with point mutations in the B-Box to stably bind transcription components. This assay has been used as an indirect method to measure the ability of TFIIIC to interact with tDNA (10, 21) and provides an advantage over the template activity assay in that measuring maxigene transcription overcomes any problems that might be incurred by RNA processing or degradation of mutant transcripts.

Table 2. Transcription and competition activities of $tDNA^{Arg}$ in HeLa cytoplasmic extract as a function of B-Box point mutations. Data are reported as percentage mean activity ± SD (N = 3 - 5). Wild type base sequences are underlined and have relative activity levels of 100%. Indicated are point mutations that significantly decreased both activities (*), only the transcription level (**), or only the competition level (***).

tDNA	Major Groove (see Table3)	Relative Transcription Level	Relative Competition Level
G52	a-a-d-h		
A	a-d-a-m	108 ± 9	100 ± 6
T	m-a-d-a	90 ± 12	94 ± 3
C***	h-d-a-a	104 ± 7	85 ± 4
G53	a-a-d-h		
A*	a-d-a-m	4 ± 5	10 ± 5
T*	m-a-d-a	23 ± 13	20 ± 8
C*	h-d-a-a	16 ± 9	15 ± 9
T54	m-a-d-a		
G*	a-a-d-h	46 ± 12	28 ± 8
A	a-d-a-m	99 ± 10	99 ± 3
C	h-d-a-a	101 ± 12	102 ± 1
T55	m-a-d-a		
G*	a-a-d-h	1 ± 2	0 ± 9
A*	a-d-a-m	34 ± 12	23 ± 1
C*	h-d-a-a	0.2 ± 0	3 ± 3
C56	h-d-a-a		
G*	a-a-d-h	8 ± 6	4 ± 5
A*	a-d-a-m	2 ± 3	1 ± 7
T*	m-a-d-a	37 ± 4	27 ± 5
G57	a-a-d-h		
A***	a-d-a-m	102 ± 14	70 ± 8
T***	m-a-d-a	91 ± 12	59 ± 9
C*	h-d-a-a	45 ± 5	27 ± 9
A58	a-d-a-m		
G*	a-a-d-h	81 ± 4	72 ± 3
T***	m-a-d-a	93 ± 7	69 ± 7
C	h-d-a-a	100 ± 9	100 ± 2
C61	h-d-a-a		
G*	a-a-d-h	35 ± 2	24 ± 9
A**	a-d-a-m	74 ± 10	89 ± 5
T*	m-a-d-a	38 ± 5	38 ± 3
C62	h-d-a-a		
G*	a-a-d-h	57 ± 2	54 ± 1
A*	a-d-a-m	68 ± 7	56 ± 3
T	m-a-d-a	91 ± 8	88 ± 4
C54 T57 C58		0.5 ± 0.5	0 ± 6
"Reverse Construct"		0 ± 0	0 ± 6

The results of the competition studies are presented in Table 1 for the *Drosophila* extract and Table 2 for the HeLa extract. A representative assay performed in the HeLa cytoplasmic extract is shown in Figure 3. Similar to their effects on transcription, point mutations at positions 53, 55, and 56, as well as the point mutation G_{54} drastically decreased the ability of $tDNA^{Arg}$ to stably bind transcription components. With a few exceptions, the results agree well with the effects that the point mutations had on template activity. The A_{52} mutation caused a significant decrease in the transcription level in the *Drosophila* extract (83%) yet did not cause a significant decrease in the relative competition level (Table 1). In addition, in *Drosophila*, C_{57} significantly reduced the competition ability without altering the relative transcription level (Table 1). Several point mutations effected a similar response in HeLa; A_{57}, T_{57}, and T_{58} each had little affect on the wild type transcription level but each decreased the competition level significantly (Table 2). The point mutation C_{52} significantly decreased template activity as well as stable factor binding in *Drosophila* but decreased only the level of competition in HeLa (Tables 1 and 2). The effect of mutations at positions 57, 58, 61, and 62 varied slightly from wild-type dependent on the substituted base pair. In HeLa extracts, C_{57} compromised template activity and stable complex formation whereas in *Drosophila*, C_{57} compromised only stable complex formation (Tables 1 and 2). In *Drosophila*, the A_{61} mutation lead to a drastic reduction in the template activity but only slightly affected stable complex formation (Table 1).

Analysis of the B-Box Specificity Determinants

There are four principal interaction sites on the major groove side of a base pair and three sites on the minor groove side (22, 24, and see Table 3). The nucleotide determinants of the B-Box were specified and analyzed systematically for the base pair combinations A·T, T·A, G·C, and C·G (Table 1). As an example of the analysis, consider the putative interaction of a factor at G_{53}. The decrease in transcriptional activity effected

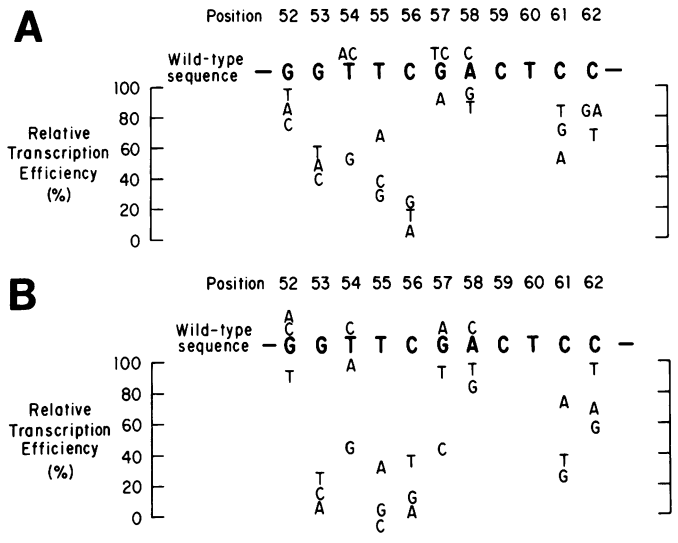

Figure 2. The effects of B-Box point mutations on transcription of a *Drosophila* $tRNA^{Arg}$ gene in (A) *Drosophila* and (B) HeLa cytoplasmic extracts.

by point mutations at position 53 indicates that G_{53} contributes to the specificity of the transcription factor interaction. The G·C base pair has a unique major groove site occupancy of a-a-d-h (see Table 3). Since no other base pair has this occupancy, the inactivity of A_{53}-, T_{53}-, or C_{53}-containing templates is readily explained by a requirement for this site occupancy. Similar arguments hold for T_{55} and C_{56}, in so far as templates containing other site configurations are essentially inactive in transcription (Tables 1 and 2). However, the chemical constituents of the A_{55} sites are furthest removed from those of T_{55} when compared to G_{55} and C_{55}, and yet A_{55} in *Drosophila* and HeLa extracts respectively effected 66% and 34% relative transcriptional levels. This result is not explained in terms of major groove contributions and suggests that base pair specificity sites may be only one of a set of possible features in the B-Box sequence that contribute to transcription factor binding.

The notion that interactions other than those provided by the bases in the major or minor grooves contribute to factor specificity is further supported by the effects that point mutations at position 54 had on transcription. T_{54}, A_{54}, and C_{54} did not compromise the wild-type transcription efficiency of $tDNA^{Arg}$ (Table 1), however, since G_{54} (a-a-d-h) drastically reduced transcription

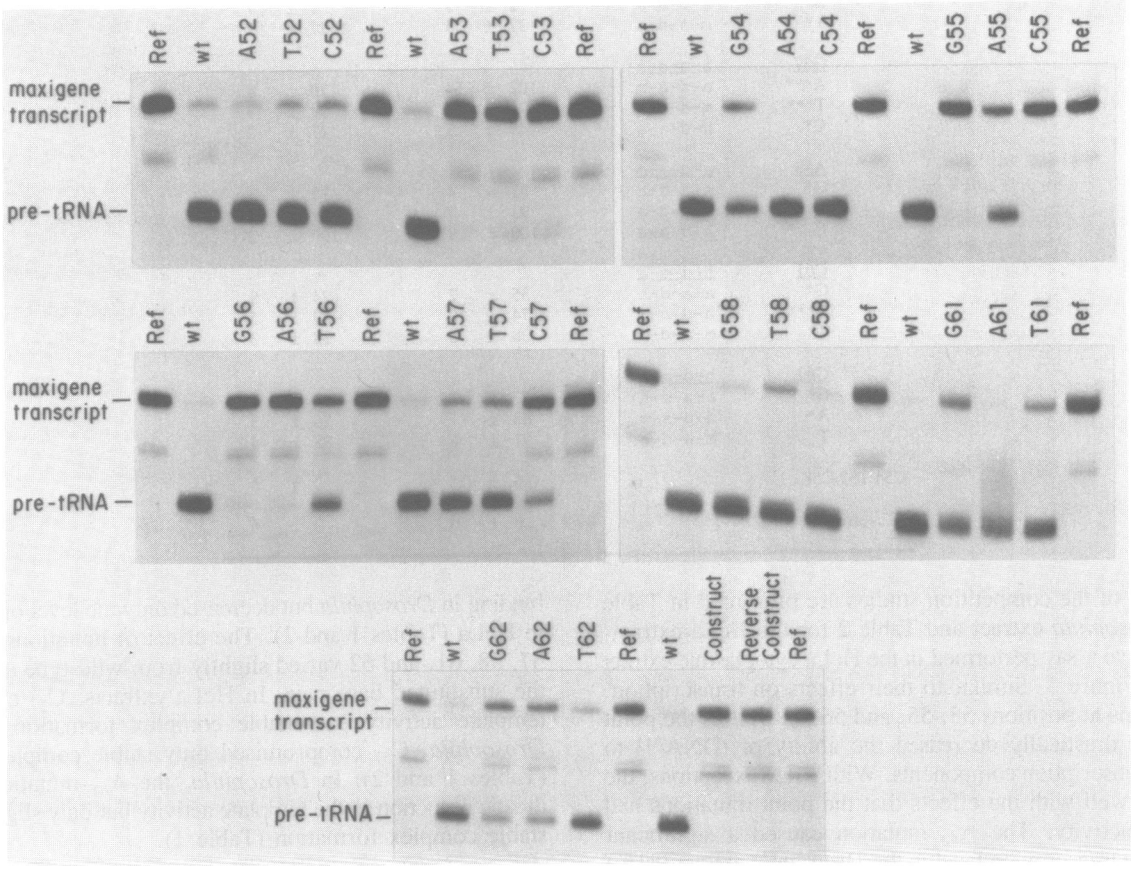

Figure 3. Effect of B-Box point mutations on the ability of a Drosophila $tRNA^{Arg}$ gene to stably sequester HeLa transcription components (Table 2). Pre-tRNA indicates the transcript produced from wild type and mutant $tDNA^{Arg}$; precursor tRNA processing activity was low in HeLa cytoplasmic extracts. This is a representative template exclusion competition assay of mutant tDNA in HeLa cytoplasmic extract, similar competition assays were performed using *Drosophila* extracts the results of which are presented in Table 1. Test DNAs are indicated above each lane. In these assays, quantitation of the maxigene transcript provided a measure of the relative ability of the test DNA to stably sequester transcription components during a preincubation period of 15 min. The relative competition levels shown in Tables 1 and 2 were determined as the amount of maxigene transcription in the presence of pUC8 (noncompeted amount) minus the amount in the presence of competitor DNA, as a percentage of, the noncompeted amount minus the amount of maxigene transcription in the presence of the wild type gene pArg3.106.

efficiency, it seems that the acceptor-acceptor (a-a) configuration of the W1 and W2 sites is not tolerated at position 54. This result was surprising since template and competition activities were not affected by the other W1, W2 site configurations and leaves unexplained the discernable feature common to the W1 and W2 sites of base pairs involving A, C and T and yet distinct from G that could explain the observed effects of these point mutations (Table 1). Therefore, if position 54 contributes to the specificity of factor interaction, this specificity might not be realized through the interaction of functional groups in the major or minor grooves.

Similarly, the effects of the point mutations at positions 52, 57, 58, 61 and 62 cannot be explained in terms of critical major groove contributions. In the HeLa transcription system however, a possible hydrogen bond acceptor at sites W1 or W2 of position 57 affords a reasonable level of transcription activity. This particular site occupation is satisfied by G·C (a-a-d-h) in this position in wild-type and is met, arguably, in the mutants by A·T (a-d-a-m) or T·A (m-a-d-a) but not by C·G (h-d-a-a, Table 2). A similar argument can be made for a primary determinant in the HeLa system being a hydrogen bond donor in the W2 site at position 61 (Table 2).

DISCUSSION

All possible point mutations at nucleotide positions 52–58, 61 and 62 of the B-Box of a *Drosophila* $tRNA^{Arg}$ gene were constructed and the resultant mutants were analyzed for their effects on template activity and stable complex formation. While this study was initiated to distinguish those nucleotides that specify B-Box ICR function from those that are conserved due to their importance in the tRNA molecule, the results support the proposal that maximum activity is afforded a tRNA gene by the presence of a B-Box sequence that satisfies the tDNA T-region consensus sequence, GG(AT)TC(GA)ANNCC (30). Although several point mutants (9 of 27 using *Drosophila* extract and 10 of 27 using HeLa extract) were found which supported wild-type transcription activity, no one position was shown to be superfluous to B-box activity. This finding suggests that future DNA-factor binding studies should also include an analysis of nucleotide positions 59 and 60 since these may not be neutral for promoter function as originally considered (1, 2).

It is well established that the B-Box ICR is involved in directing the activity of TFIIIC (reviewed in 1). However, whether the B-Box ICR also has a role in directing the activities of other tDNA transcription components is not known. In this regard then, mutations in the B-Box may, either directly or indirectly, affect the functions of TFIIIB and RNA polymerase III. The present transcription and stable complex formation assays using crude extracts are not able to discern the transcription components affected by the particular point mutations. Since the B-Box point mutations potentially affect multiple components required for tDNA transcription, caution must be exercised in drawing specific conclusions regarding DNA-factor binding.

The preference for the T-region consensus sequence suggests that all the positions of the 11-bp B-Box sequence are directly involved in factor interactions. While the wild-type $tDNA^{Arg}$ sequence, which obeys consensus, afforded maximum template activity, the B-Box ICR displayed a degree of sequence tolerance in that many of the mutant tDNAs supported transcription, albeit at levels reduced from wild-type. Since positions 53, 55, and 56 represent the only base pairs that could not tolerate change, it is likely that these nucleotides are the major determinants specifying B-Box function. The G_{56} point mutation in the yeast $tRNA^{Tyr}$ gene first identified C_{56} as a critical determinant in B-Box function (17). This identity is corroborated by the present analysis in both the *Drosophila* and HeLa transcription systems and is extended since, for $tDNA^{Arg}$, the point mutations A_{56} and T_{56} effectively lead to loss of B-Box function. The general observation that proteins maximize their available interactions with DNA (41) provides a framework to understand the interaction of a factor(s) with the B-Box ICR. It seems likely from the mutation results that a core sequence consisting of G_{53} and $T_{55}C_{56}$ initiates factor interaction with the B-Box. While the consensus sequence provides the optimal sequence in maximizing the interactions initiated by this core, other sequences can suffice but to a lesser extent.

The HeLa transcription apparatus was less tolerant of mutations in the critical positions 53, 55, and 56 than was the *Drosophila* cell extract. In addition, while the inactivity of the 'reverse construct' mutation in HeLa cell extract indicated that the B-Box ICR had been effectively removed, this DNA supported a 10% level of transcription in the *Drosophila* system. A similar observation was made in an earlier study in which deletion of the B-Box ICR resulted in mutant tDNAs that could still support a low level of transcription in the *Drosophila* system but not in HeLa cell extract (38). In view of the effects of the point mutations, this might reflect a difference in the requirements of the respective TFIIIC activities similar to the difference between yeast and human TFIIIC. Since human TFIIIC2 binds only to the B-Box ICR (15), a critical mutation precluding binding is likely to prevent transcription complex formation and transcription because the A-Box binding component, TFIIIC1, does not associate with tDNA without TFIIIC2 (15). On the other hand, if the activity of *Drosophila* TFIIIC is similar to that of yeast (6) and binds to both the A- and B-Box ICRs it could still associate weakly with the gene through the A-Box ICR in the absence of a B-Box ICR and establish a low level of transcription activity.

Except for positions 53, 55, and 56, a pattern of base pair preferences at specific positions within the B-Box was not

Table 3. Base-pair specificity sites of the B-Box sequence of $tDNA^{Arg}$ after Seeman *et al.* (22). Present on the indicated base pair at the indicated site is: a, hydrogen-bond acceptor; d, hydrogen-bond donor; m, methyl group; h, hydrogen (C-H).

		Contents of Site						
		Major Groove				Minor Groove		
Position	Base Pair	W1	W2	W2'	W1'	S1	S2	S1'
5'								
52	G•C	a	a	d	h	a	d	a
53	G•C	a	a	d	h	a	d	a
54	T•A	m	a	d	a	a	d	a
55	T•A	m	a	d	a	a	d	a
56	C•G	h	d	a	a	a	d	a
57	G•C	a	a	d	h	a	d	a
58	A•T	a	d	a	m	a	d	a
59	N							
60	N							
61	C•G	h	d	a	a	a	d	a
62	C•G	h	d	a	a	a	d	a
3'								

revealed. For example, tDNA containing A_{55} displayed a higher activity compared to tDNA containing either G_{55} or C_{55}. Since the base pairs T·A and A·T have no common sites distinct from G·C or C·G base pairs that could account for the activity of the A_{55}-containing tDNA, the possibility is raised that other structural features of the B-Box sequence are important for its function. One explanation for this result derives from an observation made for the 434 operator-repressor interaction (42) and relates to the potential of the DNA to be bent or overwound on interaction with the factor. Conceivably, a mutation from a T·A base pair to an A·T base pair would not necessarily affect this ability, whereas, a change to G·C or C·G would make bending or unwinding thermodynamically less favorable (42). Thus, as evidenced by the decrease in transcription, the A_{55} mutation probably removed a factor binding determinant but the flexibility of the DNA at this site might not have been changed allowing the factor to interact, albeit with a lessened ability. Such observations raise the possibility that the local DNA geometry is an important determinant in specifying B-Box function.

Further support for the notion that DNA geometry is important in B-Box function is provided by the effect of the G_{54} mutation. While the presence of A, C, or T at position 54 affords a functional B-Box, G_{54} was deleterious to B-Box function. This result is not readily explained in terms of potential sites on the major or minor grooves of the DNA helix (see Table 3). Indeed, the exclusion of G_{54} appears to be a more critical determinant of B-Box function than the inclusion of A, C or T. From the particular base stacking observed for the pGpGp sequence (43, 44), it is conceivable that $G_{53}G_{54}$ is disallowed because it disrupts a local DNA geometry that is important for the B-Box ICR. The effects of the synthetic B-Box (C_{54}, T_{57}, and C_{58}) also support this notion. This triple mutant, which contains only individually neutral (T_{57}) or promoter-up mutations (C_{54}, C_{58}), was transcribed in the *Drosophila* extract at only 48% of the wild-type level, and allowed stable complex formation to an even lesser extent. It was inactive in HeLa cell extract. The B-Box ICR sequence resulting from this triple point mutation consisted of a ten nucleotides long pyrimidine stretch, CTCTCCTCCT, which may have altered the DNA conformation to the extent that B-Box function was compromised.

Local variations in the structure of DNA are intrinsic to the sequence of bases (25, 41) but whether these variations contribute to the specificity of factor binding was, until recently, unknown. The finding that the *E. coli trp* repressor interacts with its cognate operator through the phosphates in the sugar-phosphate backbone strikingly domonstrates the importance that DNA geometry can have in specifying a protein-DNA interaction (25). The need for the B-Box ICR to adopt or to be able to accommodate a specific conformation in order to allow factor interactions would certainly contribute to the selective pressure on its sequence, perhaps even on nucleotides not directly involved in DNA-protein interactions.

ACKNOWLEDGEMENTS

This work was supported by grants from the American Cancer Society and University of California, Irvine, Faculty Research Funds to S. J. S. and from the Australian Research Council to T. S. S. B. A. G. is the recipient of a Faculty of Biological and Behavioral Sciences Postdoctoral Scholarship.

REFERENCES

1. Geiduschek, E. P., and Tocchini-Valentini, G. P. (1988) Ann. Rev. Biochem. **57**, 873–914.
2. Sharp, S. J., Schaack, J., Cooley, L., Johnson Burke, D., and Söll, D. (1985) CRC Crit. Rev. Biochem. **19**, 107–144.
3. Lassar, A. B., Martin, P. L., and Roeder, R. G. (1983) Science **222**, 740–748.
4. Schaack, J., Sharp, S., Dingermann, T., and Söll, D. (1983) J. Biol. Chem. **258**, 2447–2453.
5. Fuhrman, S. A., Engelke, D. R., and Geiduschek, E. P. (1984) J. Biol. Chem. **259**, 1934–1943.
6. Baker, R. E., and Hall, B. D. (1984) EMBO J. **3**, 2793–2800.
7. Johnson Burke, D., and Söll, D. (1985) J. Biol. Chem. **260**, 816–823.
8. Segall, J., Matsui, T., and Roeder, R. G. (1980) J. Biol. Chem. **255**, 11986–11991.
9. Stillman, D. J., and Geiduschek, E. P. (1984) EMBO J. **3**, 847–853.
10. Camier, S., Gabrielson O., Baker, R., and Sentenac, A. (1985) EMBO J. **4**, 491–500.
11. Marzouki, N., Camier, S., Ruet, A., Moenne, A., and Sentenac, A. (1986) Nature **323**, 176–178.
12. Gabrielson, O.S., Marzouki, N., Ruet, A., Sentenac, A., and Fromageot, P. (1989) J. Biol. Chem. **264**, 7505–7511.
13. Yoshinaga, S. K., Boulanger, P. A., and Berk, A. J. (1987) Proc. Natl. Acad. Sci. USA **84**, 3585–3589.
14. Dean, N., and Berk, A. J. (1987) Nucleic Acids Res. **15**, 9895–9907.
15. Dean, N., and Berk, A. J. (1988) Mol. Cell. Biol. **8**, 3017–3025.
16. Folk, W. R., Hofstetter, H., and Birnstiel, M. L. (1982) Nucleic Acids Res. **10**, 7153–7162.
17. Allison, D. S., Goh, S. H., and Hall, B. D. (1983) Cell **34**, 655–664.
18. Baker, R. E., Gabrielson, O., and Hall, B. D. (1986) J. Biol. Chem. **261**, 5275–5282.
19. Folk, W. R., and Hofstetter, H. (1983) Cell **33**, 585–593.
20. Traboni, C., Ciliberto, G., and Cortese, R. (1984) Cell **36**, 179–187.
21. Newman, A. J., Ogden, R. C., and Abelson, J. (1983) Cell **35**, 117–125.
22. Seeman, N. C., Rosenberg, J. M., and Rich, A. (1976) Proc. Natl. Acad. Sci. USA **73**, 804–808.
23. von Hippel, P. H., and Berg, O. G. (1986) Proc. Natl. Acad. Sci. USA **83**, 1608–1612.
24. McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P., Boyer, H. W., Grable, J., and Rosenberg, J. M. (1986) Science **234**, 1526–1541.
25. Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., and Sigler, P. B. (1988) Nature **335**, 321–329.
26. Jordan, S., and Pabo, C. O. (1988) Science **242**, 893–899.
27. Aggarwal, A. K., Rodgers, D, W., Drottar, M., Ptashne, M., and Harrison, S. C. (1988) Science **242**, 899–907.
28. Pabo, C.O., and Sauer, R. T. (1984) Ann. Rev. Biochem. **53**, 293–321.
29. Sharp, S., DeFranco, D. Dingermann, T., Farrell, P., and Söll, D. (1981) Proc. Natl. Acad. Sci. USA **78**, 6657–6661.
30. Galli, G., Hofstetter, H., and Birnstiel, M.L. (1981) Nature **294**, 626–631.
31. Schaack, J., Sharp, S., Dingermann, T., Johnson Burke, D., Cooley, L., and Söll, D. (1984) J. Biol. Chem. **259**, 1461–1467.
32. Kramer, W., and Fritz, H.J. (1987) Methods in Enzymol. **154**, 350–367.
33. Kramer, B., Kramer, W., and Fritz, H. J. (1984) Cell **38**, 879–887.
34. Miyada, C.G., and Wallace, R.B. (1987) Methods in Enzymol. **154**, 94–107.
35. Tabor S., and Richardson C.C. (1987) Proc. Nat. Acad. Sci. USA. **84**, 4767–4771.
36. Lofquist, A. K., Garcia, A. D., and Sharp, S. J. (1988) Molec. Cell. Biol. **8**, 4441–4449.
37. Stewart, T. S., Söll, D. G., and Sharp, S. J. (1985) Nucleic Acids Res. **13**, 435–447.
38. Sharp, S., Dingermann, T., Schaack, J., DeFranco, D., and Söll, D. (1983) J. Biol. Chem. **258**, 2440–2446.
39. Rich, A., and Kim, S.H. (1976) Ann. Rev. Biochem. **45**, 805–860.
40. Sprinzl, M., Hartmann, T., Weber, J., Blank, J., and Zeidler, R. (1989) Nucleic Acids Res. **17s**, 1–172.
41. Richmond, T.J. (1987) Nature **326**, 18–19.
42. Koudelka, G. B., Harrison, S. C., and Ptashne, M. (1987) Nature **326**, 886–888.
43. McCall, M., Brown, T., Hunter, W. N., and Kennard, O. (1986) Nature **322**, 661–664.
44. Rhodes, D., and Klug, A. (1986) Cell **46**, 123–132.