Molecular coevolution of mammalian ribosomal gene terminator sequences and the transcription termination factor TTF-I

(RNA polymerase I/transcription factors/species specificity/DNA-protein interactions/helix-turn-helix motif)

RAYMOND EVERS[†] AND INGRID GRUMMT

Division of Molecular Biology of the Cell II, German Cancer Research Center, 69120 Heidelberg, Germany

Communicated by M. Lindauer, Theodor-Boveri-Institut für Biowissenschaften (Biozentrum) der Universität Würzburg, Würzburg, Germany, March 24, 1995 (received for review February 6, 1995)

ABSTRACT Both the DNA elements and the nuclear factors that direct termination of ribosomal gene transcription exhibit species-specific differences. Even between mammals-e.g., human and mouse-the termination signals are not identical and the respective transcription termination factors (TTFs) which bind to the terminator sequence are not fully interchangeable. To elucidate the molecular basis for this species-specificity, we have cloned TTF-I from human and mouse cells and compared their structural and functional properties. Recombinant TTF-I exhibits species-specific DNA binding and terminates transcription both in cell-free transcription assays and in transfection experiments. Chimeric constructs of mouse TTF-I and human TTF-I reveal that the major determinant for species-specific DNA binding resides within the C terminus of TTF-I. Replacing 31 C-terminal amino acids of mouse TTF-I with the homologous human sequences relaxes the DNA-binding specificity and, as a consequence, allows the chimeric factor to bind the human terminator sequence and to specifically stop rDNA transcription.

In eukaryotes, ribosomal precursor RNA is synthesized in a complex, ordered series of events resulting in molecules with precise 5' and 3' ends. Like all other steps in RNA synthesis, these reactions are complex ones that require multiple protein-nucleic acid and protein-protein interactions. The key event in transcription termination by RNA polymerase I (pol I) is the binding of a protein, TTF-I (transcription termination factor for pol I), to a terminator element located downstream of the pre-rRNA coding region. Interestingly, species-specific differences exist between both the terminator sequence and the corresponding interacting factor. The sequence motifs that mediate pol I transcription termination in yeast and Xenopus are clearly different from those found in mammals (for review, see ref. 1). Even in closely related mammals the cis-acting sequences and trans-acting factors are not identical. In mouse and rat, the termination signal is an 18-bp sequence motif AGGTCGACCAGA/TT/ANTCCG, termed "Sal box," because it contains a Sal I restriction enzyme recognition sequence (underlined) (2). Deletion mutants and base substitutions that disrupt binding of TTF-I to this target sequence abolish pol I-dependent termination (3). The human terminator element (GGGTCGACCAG) is shorter than its rodent counterpart. It comprises only 11 nucleotides, 10 of which are identical to the proximal portion of the murine Sal box (4). In addition, the physicochemical properties and the sequence specificity of the mouse TTF-I (mTTF-I) and human TTF-I (hTTF-I) have been shown to be different. The mouse terminator element is recognized both by the murine and the human factor. In contrast, the human termination signal is recognized only by the homologous human and not by the heterologous mouse factor (4, 5), indicating that changes in terminationsignal sequences have been accompanied by a molecular coevolution of the gene encoding the termination factor.

Recently, we have isolated the cDNA encoding mTTF-I (6). Here, we describe the cloning and functional properties of the corresponding human factor and compare the primary structure and sequence specificity of mTTF-I and hTTF-I.[‡] We find that the amino acid sequence of TTF-I from both species is highly conserved within the C-terminal half, whereas the N-terminal half exhibits much-less pronounced sequence conservation. Interestingly, the species-specific differences in DNA binding reside in the conserved C-terminal region. Replacement of 31 C-terminal amino acids of mTTF-I by the corresponding hTTF-I sequence enables the chimeric factor to interact with the human termination signal. The data suggest that the C-terminal part of the DNA-binding domain is responsible for the sequence selectivity of TTF-I.

MATERIALS AND METHODS

Cloning and Expression of cDNA Encoding hTTF-I. Cloning the cDNA encoding mTTF-I has been described (6). To isolate cDNA clones encoding hTTF-I, a HeLa Lambda ZAP II cDNA library was screened under low-stringency conditions with a 1347-bp Pvu II fragment from mTTF Δ N323. Two overlapping cDNAs were isolated which encode the C-terminal 660 amino acids of hTTF-I. The 5' end of hTTF-I cDNA was cloned by a PCR-based strategy using a 5'-ampliFINDER RACE kit (Clontech). hTTF-I derivatives were inserted into either the mammalian expression vector pSG (pSG-hTTF) or the bacterial expression vector pRSET, introducing a histidine tag at the 5' end of the cDNA. The chimeric constructs containing parts of m- and hTTF-I (TTFswap 1-5) are schematically illustrated in Fig. 4A. Specific details of the cloning will be provided on request. The histidine-fusion proteins were expressed in Escherichia coli INVaF' and purified on Ni-agarose columns as described (6). The chimeric proteins were synthesized with the TNT coupled transcription/translation system (Promega)

Plasmid Constructs. The pUC9-derived minigene constructs pMrmSB and pMrhSB contain mouse rRNA-encoding DNA (rDNA) promoter sequences from -170 to +155 fused to a double-stranded oligonucleotide containing the mouse (5'-CCCGGGGATCCTTCGG<u>AGGTCGACCAGTACTCCG</u> GGCGACCCCGGGATC-3') or the human (5'-GATCTCCG-CAC<u>GGGTCGACCAG</u>CAGAATT-3') Sal box sequence. For transient transfection experiments, a 167-bp fragment from the bacterial chloramphenicol acetyltransferase (CAT) reporter gene

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: pol I, RNA polymerase I; TTF-I, transcription termination factor for pol I, mTTF-I and hTTF-I, mouse and human TTF-I, respectively; rDNA, rRNA-encoding DNA; EMSA, electrophoretic mobility-shift assay; CAT, chloramphenicol acetyltransferase.

^{*}Present address: Division of Molecular Biology, The Netherlands Cancer Institute, Plesmalaan 121, 1066CX Amsterdam, The Netherlands.

[‡]The sequences reported in this paper have been deposited in the GenBank data base [accession nos. X83973 (human TTF-I) and X83974 (mouse TTF-I)].

(Stu I–Pvu II fragment from pSV2CAT) was inserted into the Ava I site (at position +131) within the promoter fragment of pMrhSB to yield pMrChSB.

In Vitro Transcription and Mapping of 3' Ends of RNA. A 25- μ l transcription assay contained 7 μ l of S100 extract, 15 ng of template DNA linearized with Nde I, and varying amounts of recombinant TTF-I (6). RNA, either taken from a 25- μ l in vitro transcription reaction or extracted from transfected NIH 3T3 cells (10 μ g of cellular RNA) was mixed with the 3'-end-labeled probe (HindIII-Sal I fragment from pMrChSB), precipitated with ethanol, and dissolved in 20 μ l of hybridization buffer (80% formamide/0.4 M NaCl/40 mM Pipes, pH 6.4/1 mM EDTA). After hybridization for 12 h at 50°C, the reaction mixtures were diluted with 300 μ l of nuclease S1 buffer and digested with 100 units of nuclease S1 (Boehringer Mannheim) for 60 min at 30°C. The hybrids were precipitated with isopropanol, dissolved in 80% formamide, and analyzed on a 6% polyacrylamide sequencing gel.

DNA-Binding Assays. DNA binding was determined in bandshift assays as described (5, 7). The reaction mixtures (25 μ l) contained 7.5 fmol of double-stranded labeled oligonucleotides in binding buffer [12 mM Tris·HCl, pH 8.0/100 mM KCl/5 mM MgCl₂/0.1 mM EDTA/0.5 mM dithioervthritol/8% (vol/vol) glycerol] and 2 μ g poly(dI-dC). After incubation for 30 min at room temperature, protein-DNA complexes were separated by electrophoresis on nondenaturing 8% polyacrylamide gels. The sequence of the mouse Sal box oligonucleotide is 5'-CCCGGGATCCTTCGGAGGTCGACCAGTACTCCG-GGCGACCCCGGGATC-3'; the corresponding human oligonucleotide (hSB) is 5'-CCCGGGATTCCGCACGGGTCGAC-CAGCAGAATTCGCGACCCCGGGATC-3'. The mutant oligonucleotide (mSB*) contains two nucleotide exchanges in the mouse Sal box (5'-GATCCTTCGGAGC*G*CGACCAG-TACTCCGGGCGACA-3').

RESULTS

Sequence Comparison of m- and hTTF-I. Recently, we have described the molecular cloning of a cDNA encoding mTTF-I (6). For cloning the corresponding human factor, an mTTF-I probe was used to screen a human cDNA library. Overlapping clones were isolated containing a 2.6-kb open reading frame

_ _

which encodes a polypeptide of 886 amino acids. The deduced amino acid sequence of hTTF-I and its alignment with mTTF-I is displayed in Fig. 1. Significantly, the C-terminal 400 amino acids of the mouse and human factors are 80% identical and 85% similar. Comparison of this region with the protein sequence library revealed a striking homology to the DNAbinding domains of the yeast transcription factor Reb1p (8, 9) and the protooncogene c-Myb (10). We have shown that this part of mTTF-I contains both the DNA-binding and the termination domains and exerts most, if not all, TTF-I functions (6). The N-terminal half, on the other hand, is much more divergent. Homology is only observed if large gaps are introduced. A remarkable feature of this N-terminal part of TTF-I is the presence of regular stretches of positively charged amino acids whose length and position differ between hTTF-I and mTTF-I. Since, however, the DNA-binding specificity resides within the C-terminal region of TTF-I (6), the species-specific differences in DNA binding cannot be attributed to differences within the N-terminal half of TTF-I but should reside in the conserved C-terminal part.

Species-Specificity of m- and hTTF-I. To study the specificity of recombinant TTF-I, histidine-tagged versions of the human and the mouse factors were expressed in E. coli, and purified proteins were assayed for DNA binding and transcription termination. We have used mTTF Δ N323 and hTTF Δ N292, truncated versions of the murine and human factors which exhibit properties similar or even identical to cellular TTF-I (6). The DNA binding of the affinity-purified proteins was assayed in an electrophoretic mobility-shift assay (EMSA) by using oligonucleotide probes containing either the murine or the human terminator sequence (Fig. 2A). To determine the specificity of binding, the reactions were carried out in the presence of excess competitor oligonucleotides: the mutant mouse Sal box (lanes 1, 4, 7, and 10), the wild-type mouse Sal box (lanes 2, 5, 8, and 11), or the human Sal box (lanes 3, 6, 9, and 12) sequence. Clearly, both mTTF-I and hTTF-I specifically interact with their cognate target sequence. Moreover, hTTF-I binds to the mouse terminator with the same specificity as the murine factor (Fig. 2A, lanes 1-3 and 4-6). On the other hand, mTTF-I interacts only with the mouse probe (Fig. 2A, lanes 1–3) and not with the human sequence (Fig. 2A, lanes 7-9). This result is in accord with the

hTTF-I	1	MEGESSINE IHTPVSDKKKRKCEIHKERPORHSHEIFR
mTTF-I	1	MRGCTENERTHTETLYKKKRWSEVSEKRPORCPSOCLESKOPOVSVLGKRRRASOTPAQETLESEWPORAKKKKRRGEPOTPTOESIESEOPPVSLLGKR
hTTF-I mTTF-I	51 101	RRESQTPAQENSESEQPRKARRARKRESEQPTSSILKTPETFLKARKTTSAHKKKKNSVLEVDKERGINUVDKENINNTPKHERKDVDVVCVD
hTTF-I	130	MSIEGKLPRKPRIDREOVIARSHAHKSEATHSRVRERRNRKHORRAASMESORARDTLPOSESHOEESWLSVGPGGEITELPASAHKNKSKKKKKSSNR
mTTF-I	196	MSKGGRSAKVRHIGHTPA ARPOHEGCRHILGDVRSRRKCKHLOKMAHMDV
hTTF-I	230	EYETLAMPEGSQAGREAGTDMORSOPTVGLDDETPQLLGPTHKKKSKKKKKKKSNHQEFRALAMFREGSQVGSFVGADMQESRPAVOUHOBTAGIFAPAYK
mTTF-I	246	
hTTF-I	330	NKSKKKKKENHORFRAVAMPESIESAYPEGSOVGE.EVGIVEGSTALKGKKESMEIKKKSKKKKLTEVKRARVGGDEGVPSKNSESTIFDSWEGDGAM
mTTF-I	280	RSIKKNVFRG. ORTHPI. PDSIDDSETISERLDGTHHGGAVGAGECESTKESHEIKKKSKKKKKKKKKKSVALA. ISSDSASVTDSKAKNALVDSBEGSGAV
hTTF-I	429	MEEGVKSRIRQKKTQACLASKHVQEAFRLEFANEEH. NVITTAEDEE IRYLEADSGDADDSDADLGSAVROLCEFI FNIIKLRATST INRMYRDDI ERIKEF
mTTF-I	375	REEDVDHRFAEABAQACSTEKHREAMQRLEHTHEBESNEISASNEJARH ISHDRRESDDSDVDLGSAVROLREFI FDI DERAATTIIRRMYRDDI GLVKEF
hTTF-I	528	KAQGVAIINFGKFSVKENKCIEKNVEDFIALTGIESADKLLYTDRYPEERSVIITNLKRYSFRLHIGRNIARPWKLIYYRAKNFDVNNYKGRYSFROTIER
mTTF-I	475	KAQGVDIRFGKFSPKENKCIEKNVCDFISLIGIESADKLLYTDRYPEERTIIINLKRKHAFRLHIGKGIARPWKLVYYRAKNIFDVNNYKGRYAFROTIER
hTTF-I	628	LRMYHSILGNDWRTIGHMVARHSLSVALKFSQISSORNHGAWSRSETIRELIKAVEEVILKKMSPQEIRHVDSKLQHNPHSCLSIVREKLYKGISWVEVEA
mTTF-I	575	LRHYHSIHGNDWRTIGHMVARSSLSVALKFSQIGGTRNGGAWSRHETORLIKAVEDVILKKMSPQEIRHUDSKLQHPHGRLSIVREKLYKGISWVEVEA
hTTF-I	728	KVQTRNWMQCKSKWTEILTKRMTNGRRITYGMNAIRARVSLIERLYHINVHDINEIDWEDIASAIGDVPHSYVQIKHSHLKAVYVPFWQKKTFPEIIDYI
mTTF-I	675	RVHTRNWMQCKSKWTEILTKRMTHGGFVYHGVNALQANITLIERLYHINVNDANEIDWEDICSAIGDVPHPFVQAKHYRLKAACVPFWQKKTFPEIIDYI
hTTF-I	828	YETTILPLLKEKTERMMEKRGTKIQTPAAPKCVEEREDIEYYEDDSEGGGHRKRKRRGIP · 886
mTTF-I	775	YKNSLPLLKEKTIRKMEKROGGLQTPAAPKCDEELERDIELERDIEHCDDDSDEGSPEEPSASDVQ · 833

FIG. 1. Comparison of the deduced amino acid sequence of h- and mTTF-I. Identical amino acid residues are boxed; gaps introduced to improve alignment are marked by dots.



FIG. 2. mTTF-I exhibits a more stringent sequence requirement than hTTF-I for DNA binding. (A) EMSA of recombinant mTTF Δ N323 and hTTF Δ N292 on mouse (lanes 1-6) or human Sal-box (lanes 7-12) oligonucleotides. Each binding assay was performed in the presence of a 500-fold molar excess of mutant mouse Sal-box oligonucleotide (m*), wild-type mouse Sal-box oligonucleotide (m), or human Sal-box oligonucleotide (h). (B) In vitro transcription. A total of 7 μ l of mouse S100 extract was used to transcribe the mouse minigene construct pMrmSB/Nde I (lanes 1-3) or pMrhSB/ Nde I (lanes 4-6), in the absence of recombinant TTF-I (lanes 1 and 3) or after complementation with 5 ng of either mTTF Δ N323 (lanes 2 and 5) or hTTF Δ N292 (lanes 3 and 6). The positions of read-through (RT) and terminated transcripts (term.) are marked.

previously observed species specificity of TTF-I–DNA interactions, which indicated that the murine factor exhibits a more stringent sequence recognition than does its human counterpart (4, 5).

We then studied whether hTTF-I would terminate pol I transcription *in vitro* when bound to either murine or human terminator sequence. For this, ribosomal minigene constructs were used which contain either the murine (pMrmSB) or the human (pMrhSB) Sal box fused to the mouse rDNA promoter. Again, both the murine and the human factors support termination on pMrmSB, the template containing the mouse terminator sequence (Fig. 2B, lanes 1–3). However, on the heterologous construct (pMrhSB), the murine factor fails to terminate transcription (Fig. 2B, lane 5), whereas the human factor is active (Fig. 2B, lane 6).

hTTF-I Terminates pol I Transcription *in Vivo.* Next, we tested the specificity of binding and termination in transient transfection assays. NIH 3T3 cells were cotransfected with an expression plasmid encoding truncated hTTF-I (pSG-hTTF) and a pol I-specific reporter construct (pMrChSB). The expression of recombinant hTTF-I was monitored on Western blots (Fig. 3A) and by EMSAs (Fig. 3B). Lysates from transfected cells efficiently shift the electrophoretic mobility of the human Sal box oligonucleotide (Fig. 3B, lane 2). Mock-transfected cells, on the other hand, do not exhibit specific binding activity (Fig. 3B, lane 1).

To investigate whether the truncated protein is able to perform all functions associated with pol I transcription termination in vivo-i.e., stopping the elongation reaction and release of the transcripts-the 3' ends of transcripts synthesized from the reporter construct were analyzed (Fig. 3C). The reporter represents a fusion between the mouse rDNA promoter and the human terminator. To distinguish transcripts from the reporter plasmid above the background of cellular pre-rRNA, a fragment from the CAT gene was positioned between the promoter and terminator to allow analysis of CAT-specific transcripts. The reporter plasmid was introduced into NIH 3T3 cells, together with pSG-hTTF, and transcripts terminated at the human Sal box were monitored by nuclease S1 mapping (Fig. 3C). Transcripts which were synthesized in vitro are shown for comparison (Fig. 3C, lanes 3-5). Consistent with the cross-species DNA-binding specificity of TTF-I, no



FIG. 3. Recombinant hTTF-I terminates transcription in vivo. (A) Western blot of hTTF-I expressed in NIH 3T3 cells. NIH 3T3 cells were transfected with 10 μ g of pMrChSB, together with 5 μ g of either the vector pSG (lane 1) or pSG-hTTF (lane 2), and hTTF-I expression was monitored on immunoblots by using anti-TTF-I antiserum. The position of endogenous mTTF-I is marked (p130). (B) EMSA of hTTF in transfected cells. Extracts (10 μ g protein) from cells transfected with pSG (lane 1), pSG-hTTF (lane 2), or 10 ng of bacterially expressed hTTF Δ N292 (lane 3) were assayed for binding to the hSB probe. (C) Nuclease S1 analysis of RNA synthesized from the chimeric minigene pMrChSB. A diagram of pMrChSB is shown above. The thin bar represents 5'-terminal rDNA sequences (from -170 to +155) in which a 167-bp reporter fragment (Stu I-Pvu II fragment from pSV2CAT) is inserted. The restriction sites used for probe preparation are shown. The stippled box represents the 11-bp human terminator sequence. Lanes 1 and 2, NIH 3T3 cells were transfected with 10 μ g of pMrChSB, together with 2.5 μ g of either pSG (lane 1) or pSG-hTTF (lane 2), and cellular RNA was analyzed by nuclease S1 mapping. Lanes 3-5, 15 ng of pMrChSB/Nde I was transcribed in vitro for 10 min in the absence of TTF-I (lane 3) or in the presence of 10 ng of either mTTF Δ N323 (lane 4) or hTTF Δ N292 (lane 5). The DNA probe and the protected fragments which correspond to terminated (175 nt) and processed (165 nt) transcripts are marked.

transcripts terminating at the human Sal box element could be detected in mock-transfected mouse cells (Fig. 3C, lane 1) or in in vitro reactions lacking TTF-I (Fig. 3C, lane 3). Significantly, RNA from cells transfected with pSG-hTTF yields two protected bands of 175 and 165 nt which correspond to transcripts whose 3' ends map 11 and 21 bp upstream of the human Sal box (Fig. 3C, lane 2). We have shown previously that the longer transcripts correspond to RNA molecules terminated 11 bp upstream of the TTF-I binding site, whereas the shorter ones represent 3' terminally processed molecules (11). The same ends are produced in vitro in the presence of hTTF Δ N292 (Fig. 3C, lane 5). The finding that expression of N-terminally truncated hTTF-I governs correct 3'-end formation of pol I transcripts at the human Sal box demonstrates (i) that N-terminal sequences of TTF-I do not appear to be required for termination in vivo and (ii) that the 11-bp human Sal box sequence is sufficient to mediate both termination and processing of transcripts.

C-Terminal Amino Acids of TTF-I Determine Species-Specific DNA-Binding. To investigate which differences in the primary structures of mTTF-I and hTTF-I account for the species-specific differences in DNA recognition, a series of chimeric TTF-I molecules was constructed in which defined regions of the DNA-binding domain of mTTF-I were replaced by the corresponding human TTF-I sequences (Fig. 4A). The recombinant proteins were expressed by *in vitro* translation, and equal amounts of TTF-I were assayed for binding to the mouse and human Sal box oligonucleotide (Fig. 4B). As expected, mTTF Δ N445 does not bind to the heterologous probe (compare Fig. 4B, lanes 1 and 8). Significantly, the individual swap mutants show different specificities. Exchang-



1 2 3 4 5 6 7 8 9 10 11 12 1314

FIG. 4. Species specificity of TTF-I resides in the C-terminal part of the DNA-binding domain. (A) Diagram of chimeric mouse and human TTF-I proteins. The open bars represent mTTF-I sequences, the stippled bars represent hTTF-I sequences. The hatched boxes in mTTF Δ N445 mark the regions of homology to domains I and II of c-Myb (9). The numbers indicate the first or last amino acid residue of the respective protein sequence. (B) EMSA. Equal amounts of *in vitro* translated proteins were assayed for binding to the mouse (lanes 1–7) or human (lanes 8–14) Sal box oligonucleotide.

ing the last 265 C-terminal amino acids of mTTF-I with hTTF-I sequences (TTFswap1) relaxes the binding specificity, as shown by its ability to interact with both the murine and the human DNA probe (Fig. 4B, lanes 3 and 10). On the other hand, TTFswap2, the chimeric factor in which mTTF-I amino acids 588–710 are replaced by the corresponding human sequence, hardly recognizes the human probe (Fig. 4B, lanes 4 and 11). This result, together with the observation that exchanging 175 amino acids from the murine C terminus by hTTF-I sequences (TTFswap3) relaxes mTTF-I sequence selectivity (Fig. 4B, lanes 5 and 12), suggests that differences in the very C-terminal part of TTF-I determine the binding specificity of the murine and human factors.

This suggestion is supported by a hTTF-I deletion mutant which lacks 31 C-terminal amino acids (hTTF479-854). Interestingly, this human mutant binds to the mouse probe but fails to interact with the human probe (Fig. 4B, lanes 2 and 9). Apparently, these deleted amino acids are required to bind to the human Sal box element. If this view is correct, then exchanging the last 31 amino acids of mTTF-I with the corresponding human residues should enable the mouse factor to interact with the human sequence. This is indeed the case. Both TTFswap4 and TTFswap5 gene fusions, in which the C terminus of TTFswap2 and mTTF Δ N445 was replaced by the hTTF-I terminus, bind to the human probe (Fig. 4B, lanes 6, 7, 13, and 14). However, the efficiency of binding is much weaker than that of hTTF-I. Therefore, although replacement of the 31 C-terminal amino acids of mouse TTF-I by human sequences relaxes its stringent DNA-binding activity, it is not sufficient to convert mTTF-I into an efficient human termination factor.

DISCUSSION

Most of our knowledge of pol I termination comes from work on three experimental systems: mammals (mouse and human), frog (Xenopus laevis and Xenopus borealis), and yeast (reviewed in ref. 1). The terminator sequences from mouse, frog, and yeast are different, and it appears that the proteins which bind to the individual terminators are also quite distinct. Despite these differences in both the terminator elements and the respective interacting factors, the mechanism of termination is probably similar or even identical in all species so far analyzed. All characterized pol I terminator elements display strong orientation and factor dependence, and RNA 3'-end formation occurs 11–21 nt upstream of the recognition sequence. In an attempt to gain insight into evolutionary changes between terminator sequences and their interacting factors, we have cloned h- and mTTF-I and compared their DNA-binding specificity.

Previous studies have shown that the human factor exhibits a more relaxed DNA-binding specificity than its murine counterpart-i.e., hTTF-I recognizes the proximal 11 nt of the murine termination signal, whereas the mouse factor requires the whole 18-bp mouse Sal box and, therefore, fails to interact with the shorter human terminator (4, 5). Consistent with the similarity of the human and mouse termination signal, we show here that the primary structure of hTTF-I is similar to but not identical with mTTF-I. The highest degree of sequence conservation is observed within the C-terminal half of TTF-I. This part of TTF-I is involved in DNA-binding and transcription termination and shows striking homology to the DNA-binding domain of the protooncogene c-Myb (10) and the yeast factor Reb1p (8, 9). Reb1p bound to its target site within the rDNA enhancer stops pol I transcription and therefore represents the yeast equivalent of mammalian TTF-I (12, 13). Both the homology of the DNA-binding domain of TTF-I and Reb1p and the more stringent binding specificity of the mouse compared with the human factor suggest that the terminator sequence and the recognition domain of TTF-I evolved together to permit these molecules to interlock. The availability of recombinant h- and mTTF-I will facilitate structural studies on TTF-I-DNA interactions, as well as a detailed genetic analysis of the regions of TTF-I which determine the binding specificity.

In an attempt to pinpoint the domain of TTF-I responsible for species-specific DNA binding, we have produced a series of chimeric proteins which represent fusions of the C-terminal DNA-binding region of m- and hTTF-I. The results demonstrate that the very C terminus of TTF-I plays an important role in the specificity of DNA binding. Replacement of the last 31 amino acids of mTTF-I by the corresponding human sequences relaxes the binding specificity and enables the chimeric TTF-I to recognize the human terminator element. Significantly, this part of hTTF-I and mTTF-I is much less conserved than the rest of the DNA-binding domain, having 19 amino acid exchanges. Nevertheless, it is noteworthy that the efficiency of binding of the swap4 mutant is much weaker than that of hTTF-I. Exchanging a larger region of mTTF-I by human sequences drastically increases the interaction with the human terminator. Thus, although replacement of the 31 C-terminal amino acids of mouse TTF-I by human sequences relaxes its stringent DNA-binding activity, other nonconserved amino acids within the C-terminal half of the DNA-binding domain also appear to contribute to differences in DNAprotein interactions.

Since we failed to express full-length recombinant TTF-I in E. coli, the functional studies-i.e., DNA binding and transcription termination-have been performed with truncated versions of m- and hTTF-I. These forms of TTF-I lack 322 and 291 N-terminal amino acid residues, respectively, but efficiently terminate pol I transcription (6). Cellular full-length mTTF-I (p130) binds weakly to DNA, which suggests that the presence of the N terminus inhibits DNA-protein interactions (7). It is possible that the stretches of positively charged residues within the N-terminal region may be important for natural proteolytic processing events which in turn may be required to unmask the DNA-binding domain. This type of processing could serve a regulatory function for TTF-I, but this remains to be investigated.

Recently, a model for eukaryotic transcription termination by pol I has been proposed (13). According to this model, termination occurs when a DNA-bound protein induces polymerase to pause in the context of a release element. Using an in vitro system containing only purified yeast pol I and recombinant Reb1p as the protein components, Lang et al. (13) have shown that sequences upstream of the Reb1p-binding site are required for pol I termination. By contrast, our in vivo studies (Fig. 3) did not reveal any requirement for flanking sequences in the termination process-i.e., stopping the elongation reaction and transcript release. Recombinant hTTF-I expressed in NIH 3T3 cells terminates transcription of a reporter construct which contains the mouse rDNA promoter fused to an oligonucleotide containing the 11-bp human Sal box sequence. This result is in accord with previous in vitro and in vivo studies showing that correct 3'-end formation of transcripts occurs in the absence of any flanking sequences (11, 14). Moreover, a significant portion of transcripts had undergone a 3'-terminal processing event and therefore must have been released from the template. Whether the requirement of a release element in the yeast pol I terminator reflects

species-specific differences in the mechanism of transcription termination or is due to the different experimental systems used remains to be investigated.

We are grateful to P. Chambon for providing the human cDNA library. We thank Stephen Mason for critically reading the manuscript and the other group members for support. This work was supported in part by the Deutsche Forschungsgemeinschaft (Leibniz-Programm and SFB 229) and the Fonds der Chemischen Industrie.

- 1.
- Reeder, R. H. & Lang, W. (1994) Mol. Microbiol. 12, 11-15. Grummt, I., Maier, U., Öhrlein, A., Hassouna, N. & Bachelerrie, 2. J.-P. (1985) Cell 43, 801-810.
- Grummt, I., Rosenbauer, H., Niedermeyer, I., Maier, U. & 3. Öhrlein, A. (1986) Cell 45, 837-846.
- 4. Bartsch, I., Schoneberg, C. & Grummt, I. (1988) Mol. Cell. Biol. 8, 3891-3897.
- 5. Pfleiderer, C., Smid, A., Bartsch, I. & Grummt, I. (1990) Nucleic Acids Res. 18, 4727-4736.
- Evers, R., Smid, A., Rudloff, U., Lottspeich, V. F. & Grummt, I. (1995) EMBO J. 14, 1248-1256.
- 7. Smid, A., Finsterer, M. & Grummt, I. (1992) J. Mol. Biol. 227, 635-647.
- 8. Ju, Q., Morrow, B. E. & Warner, J. R. (1990) Mol. Cell. Biol. 10, 5226-5234
- 9. Morrow, B. E., Ju, Q. & Warner, J. R. (1993) Mol. Cell. Biol. 13, 1173-1182.
- 10. Kanei-Ishii, C., Sarai, A., Sawazaki, T., Nakagoshi, H., He, D. N., Ogata, K., Nishimura, Y. & Ishii, S. (1990) J. Biol. Chem. 265, 19990-19995.
- 11. Kuhn, A. & Grummt, I. (1989) Genes Dev. 3, 224-231.
- 12. Lang, W. H. & Reeder, R. H. (1993) Mol. Cell. Biol. 13, 649-658.
- Lang, W. H., Morrow, B. E., Ju, Q., Warner, J. R. & Reeder, 13. R. H. (1994) Cell 79, 1-20.
- 14. Kuhn, A., Normann, A., Bartsch, I. & Grummt, I. (1988) EMBO J. 7, 1497-1502.