11. J. J. Neefjes, F. Momburg, G. J. Hämmerling, *Science* **261**, 769 (1993); J. C. Shepherd *et al.*, *Cell* **74**, 577 (1993); M. J. Androlewicz, K. S. Anderson, P. Cresswell, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9130 (1993).
12. M.-T. Heemels, T. N. M. Schumacher, K. Wonigeit, H. L. Ploegh, *Science* **262**, 2059 (1993); F. Momburg *et al.*, *Nature* **367**, 648 (1994); T. N. M. Schumacher *et al.*, *J. Exp. Med.* **179**, 533 (1994).
13. F. Momburg, J. J. Neefjes, G. J. Hämmerling, *Curr. Opin. Immunol.* **6**, 32 (1994).
14. F. Lévy, R. Gabathuler, R. Larsson, S. Kvist, *Cell* **67**, 265 (1991).
15. H. R. Gaskins, J. J. Monaco, E. H. Leiter, *Science* **256**, 1826 (1992).
16. Y. Yang *et al.*, *J. Biol. Chem.* **267**, 11669 (1992).
17. D. B. Williams, B. H. Barber, R. A. Flavell, H. Allen, *J. Immunol.* **142**, 2796 (1989).
18. E. Degen and D. B. Williams, *J. Cell Biol.* **112**, 1099 (1991).
19. H.-G. Ljunggren *et al.*, *Nature* **346**, 476 (1990).
20. M. Matsumura, Y. Saito, M. R. Jackson, E. S. Song, P. A. Peterson, *J. Biol. Chem.* **267**, 23589 (1992).
21. E. R. Christinck, M. A. Luscher, B. H. Barber, D. B. Williams, *Nature* **352**, 67 (1991); T. Elliott, J. Elvin, V. Cerundolo, H. Allen, A. Townsend, *Eur. J. Immunol.* **22**, 2085 (1992).
22. R. G. Urban *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1534 (1994).
23. V. Cerundolo *et al.*, *Eur. J. Immunol.* **21**, 2069 (1991).
24. K. Anderson *et al.*, *J. Exp. Med.* **174**, 489 (1991).
25. R. A. Henderson *et al.*, *Science* **255**, 1264 (1992); M. L. Wei and P. Cresswell, *Nature* **356**, 443 (1992).
26. D. B. Burshtyn and B. H. Barber, *J. Immunol.* **151**, 3082 (1993); J. J. Neefjes, G. J. Hämmerling, F. Momburg, *J. Exp. Med.* **178**, 1971 (1993); M.-J. Bijlmakers and H. L. Ploegh, *Curr. Opin. Immunol.* **5**, 21 (1993).
27. T. Elliott, V. Cerundolo, A. Townsend, *Eur. J. Immunol.* **22**, 3121 (1992).
28. M. R. Jackson, M. F. Cohen-Doyle, P. A. Peterson, D. B. Williams, *Science* **263**, 384 (1994).
29. M. H. Smith, J. R. M. Parker, R. S. Hodges, B. H. Barber, *Mol. Immunol.* **23**, 1077 (1986).
30. B. Jones and C. A. Janeway, *Nature* **292**, 547 (1981).
31. K. Ozato and D. H. Sachs, *J. Immunol.* **126**, 317 (1981).
32. H. Lemke, G. J. Hämmerling, U. Hämmerling, *Immunol. Rev.* **47**, 175 (1979).
33. K. Ozato, T. H. Hansen, D. H. Sachs, *J. Immunol.* **125**, 2473 (1980).
34. EL4, MDAY-D2, RDM-4 (18), and transfected R1E cells (17) were cultured in RPMI 1640 medium containing 10% fetal bovine serum and antibiotics. For metabolic radiolabeling, cells were incubated at 37°C in methionine-free RPMI 1640 medium for 30 min and then suspended at 5×10^7 cells/ml in Met-free medium containing 0.5 mCi/ml [^{35}S]Met (>800 Ci/mmol, Amersham). After radiolabeling, aliquots of cells (3×10^6 to 1×10^7) were either lysed directly, or, in the case of pulse-chase experiments, 10 volumes of RPMI 1640 medium containing 1 mM methionine and 10% fetal bovine serum were added and aliquots of 3×10^6 cells were withdrawn after various periods of time and lysed. For lysis, cells were suspended on ice in 0.5 ml of digitonin lysis buffer (0.5% digitonin, 25 mM Hepes, pH 7.2, 10 mM $CaCl_2$, 10 mM iodoacetamide, 1% aprotinin, and 0.25 mM phenylmethylsulfonylfluoride) for 30 min. Lysates were precleared for 30 min with 1% fixed *Staphylococcus aureus* bacteria and, after centrifugation to remove cell debris and the bacteria, they were subjected to immunoprecipitation for 2 to 16 hours on ice with the following antibodies: For TAP1 and TAP2, rabbit antisera to the COOH-terminal 280 amino acids of TAP1 or the COOH-terminal 273 amino acids of TAP2 (*16*); for free or β_2M-associated K^b, K^d, or K^k heavy chains, rabbit antiserum to the peptide encoded by exon 8 of the K^b gene (anti-peptide 8) (*29*); for β_2M-associated K^b heavy chains, mAb Y-3 (*30*); for free or β_2M-associated D^b heavy chains, mAb 28-14-8s (*31*); for β_2M-associated D^b heavy chains, mAb B22-249.R1 (*32*); and for β_2M-associated L^d heavy chains, mAb 30-5-7s (*33*). Immune complexes were recovered by shaking with protein A–agarose beads for 1 hour and then the beads were washed five times in 0.1% digitonin, 25 mM Hepes, pH 7.2, 10 mM $CaCl_2$. The immune complexes were either eluted directly with SDS-PAGE sample buffer or were first treated with endo H (*28*). Samples were analyzed by SDS-PAGE and visualized by fluorography.
35. A. Velasco *et al.*, *J. Cell Biol.* **122**, 39 (1993).
36. For sequential immunoprecipitation, cells were radiolabeled with [^{35}S]Met for 15 min, lysed in digitonin lysis buffer and subjected to a first round of immunoprecipitation (*34*). Immune complexes collected on protein A–agarose were then disrupted by incubating in 0.5 ml of 0.2% SDS in water at 37°C for 1 hour and then samples were adjusted to contain 2% NP-40, 5% skim milk powder, 1 mM tris, pH 7.4, 15 mM NaCl, and 0.1 mM EDTA. Antibody (anti-peptide 8 or 28-14-8s) was then added for a second round of immunoprecipitation as described except that the protein A–agarose beads were washed only once with 1% NP40, 10 mM tris, pH 7.4, 150 mM NaCl, and 1 mM EDTA.
37. V. Cerundolo, T. Elliott, J. Elvin, J. Bastin, A. Townsend, *Eur. J. Immunol.* **22**, 2243 (1992).
38. H. Plutner, H. W. Davidson, J. Saraste, W. E. Balch, *J. Cell Biol.* **119**, 1097 (1992).
39. The Flu NP Y367-374 peptide corresponds to residues 367 to 374 of the influenza nucleoprotein plus a tyrosine at the NH_2-terminus (sequence YSNENMETM) and the VSV G 52-59 peptide corresponds to residues 52-59 of the vesicular stomatitis virus G protein (sequence RGYVYQGL) (*2*). Peptides were purchased from the Alberta Peptide Institute and were judged to be more than 95% pure by reversed-phase high-performance liquid chromatography analysis.
40. We thank S. Pind for suggesting the use of digiftonin for permeabilizing cells, B. Barber for antiserum to peptide 8, and S. Pind and Y. Yang for helpful comments. Supported by the National Cancer Institute of Canada with funds from the Canadian Cancer Society (D.B.W.).

12 April 1994; accepted 9 May 1994

The TATA-Binding Protein: A General Transcription Factor in Eukaryotes and Archaebacteria

Tracey Rowlands, Peter Baumann, Stephen P. Jackson*

The TATA-binding protein TBP appears to be essential for all transcription in eukaryotic cell nuclei, which suggests that its function was established early in evolution. Archaebacteria constitute a kingdom of organisms distinct from eukaryotes and eubacteria. Archaebacterial gene regulatory sequences often map to TATA box–like motifs. Here it is shown that the archaebacterium *Pyrococcus woesei* expresses a protein with structural and functional similarity to eukaryotic TBP molecules. This suggests that TBP's role in transcription was established before the archaebacterial and eukaryotic lineages diverged and that the transcription systems of archaebacteria and eukaryotes are fundamentally homologous.

The TATA box–binding protein TBP is an essential transcription factor for eukaryotic RNA polymerases I, II, and III, and thus appears to be required for all nuclear transcription (*1*). Binding of TBP to DNA is usually the first step in the assembly of an RNA polymerase II transcription complex. Once bound to DNA, TBP nucleates the assembly of other general transcription factors and directs the start of transcription to a site approximately 25 base pairs (bp) downstream from the TATA box. TBP also functions in transcriptional regulation because it is the target of many transcriptional activators and repressors (*2*, *3*). Reflecting its importance, TBP is exceedingly conserved throughout evolution; its core domain of 180 amino acid residues is over 80% identical between yeast and humans (*4*).

Sequence analysis shows that the three eukaryotic RNA polymerases are evolutionarily related (*5*), which indicates that they arose from a common origin. One might therefore speculate that the ancestor of eukaryotes had a single RNA polymerase and a single set of general transcription factors, one of which would have been TBP. Although the single RNA polymerase of eubacteria is related to the eukaryotic enzymes, eubacteria appear to lack homologs of TBP and of other eukaryotic general factors. The archaebacteria (archaea) constitute a third major kingdom of life (*6*, *7*). Although archaebacteria lack nuclei, resemble eubacteria in morphology, and have a single RNA polymerase, they are at least as distant evolutionarily from *Escherichia coli* as they are from humans. Archaebacterial RNA polymerases more closely resemble those of eukaryotes than those of eubacteria (*7*), and putative archaebacterial homologs of transcription factors TFIIB and TFIIS have been identified (*8*, *9*). Because many archaebacterial gene regulatory sequences map to TATA box–like motifs approximately 25 bp upstream of transcription start sites and fractionation of archaebacterial extracts has identified a transcription factor that operates through

Wellcome/CRC Institute, Tennis Court Road, and Department of Zoology, Cambridge University, Cambridge CB2 1QR, UK.

*To whom correspondence should be addressed.

these sequences (*10*), we searched for TBP-like proteins in archaebacteria.

Eukaryotic TBP molecules contain a relatively nonconserved NH_2-terminal region and a highly conserved COOH-terminal domain that consists of two imperfect direct repeats and an intervening basic region (*4*). Whereas the NH_2-terminal portion is dispensable for basal transcription, the COOH-terminal domain is essential and specifies TATA box binding and interaction with other general and regulatory transcription factors (*1–3, 11*). Within the COOH-terminal domain, several regions are essentially invariant in sequence in all TBP molecules (*4*). We synthesized degenerate oligonucleotides corresponding to these motifs and used them in polymerase chain reactions (PCRs) with genomic DNA of various archaebacteria. For the hyperthermophilic archaebacterium *Pyrococcus woesei*, one oligonucleotide pair yielded a product whose sequence was compatible with that of a TBP gene homolog. In Southern (DNA) blot analysis of restriction enzyme–digested *P. woesei* genomic DNA, the cloned PCR product hybridized to a single DNA fragment in each case, which demonstrates that it is derived from *P. woesei* and that it corresponds to a single-copy locus (Fig. 1A).

By screening *P. woesei* DNA libraries with the cloned PCR product, we isolated clones spanning its genomic locus (Fig. 1B). Sequencing showed the presence of an open reading frame that was capable of encoding a polypeptide of 21,310 daltons with homology to the conserved COOH-terminal domain of eukaryotic TBP molecules (Fig. 1B). In contrast to eukaryotic TBPs, however, the *Pyrococcus* TBP sequence has essentially no NH_2-terminal extension and is, in fact, shorter than any eukaryotic TBP reported to date.

We identified a monoclonal antibody that recognizes *Pyrococcus* TBP by overexpressing this protein in *E. coli* and screening with various antibodies raised against eukaryotic TBPs (Fig. 2A). When this antibody was used to probe a protein immunoblot of a crude *Pyrococcus* extract, a single band of approximately 20 kD was detected (Fig. 2A). This suggests that the TBP homolog gene we have cloned is expressed in *P. woesei*.

In spite of the large evolutionary distance between eukaryotes and archaebacteria, the COOH-terminal domains of *P. woesei* and of human TBPs are approximately 40% identical. The *P. woesei* factor is approximately equally related (36 to 41% overall identity) to all eukaryotic TBPs with demonstrated transcriptional properties. Furthermore, hydropathy and secondary structure predictions for *P. woesei* and for eukaryotic TBPs are very similar, and

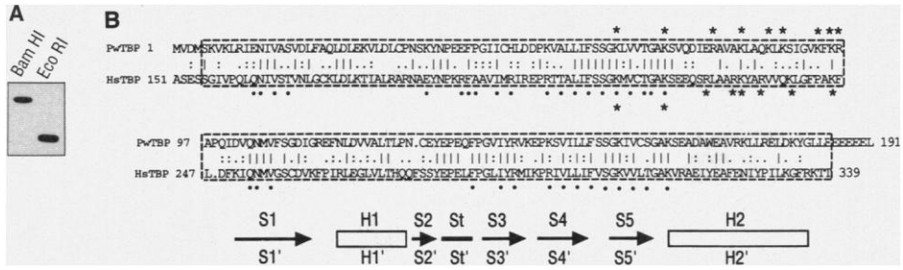

Fig. 1. (**A**) The cloned TBP-related sequence corresponds to a single-copy *P. woesei* gene. Southern blots of Bam HI– and Eco RI–digested *P. woesei* genomic DNA were probed with radiolabeled cloned PCR product and washed at 50°C in 0.5×SST and 0.1% SDS. (**B**) Comparison of the predicted *P. woesei* TBP (PwTBP) sequence with the COOH-terminal core domain of human TBP (HsTBP) (*20*). The two imperfect direct repeats are boxed, TBP residues of *A. thaliana* or *S. cerevisiae* (or both) that mediate DNA contacts are indicated by dots, and positively charged residues around the basic repeat region are indicated by asterisks. The approximate positions of the secondary structural elements defined for eukaryotic TBPs (*12, 13*) are indicated by S for strand, H for helix, and St for stirrup. The PCR conditions for cloning archaebacterial TBP homologs were as follows: (1) 94°C for 3.5 min; (2) 94°C for 1.5 min; (3) increment 50°C for 1.5 min, −1°C cycle ("touchdown"; 50° to 40°C); (4) slope 1°C/6 s to 72°C; (5) 72°C for 3 min; (6) cycle to (2) 10 times; (7) 94°C for 1.5 min; (8) 40°C for 1.5 min; (9) slope 1°C/6 s to 72°C; (10) 72°C for 3 min; and (11) cycle to (7) 25 times. Reactions containing *P. woesei* DNA and primers 8 and 10 [corresponding to amino acids YEPELF(P/A)G and FVSGK(I/V)(I/V)(I/L)TGA] yielded a product that was cloned into pBSKS+.

computer modeling indicates that the *P. woesei* sequence is entirely compatible with it assuming an equivalent structure to eukaryotic TBP molecules (Fig. 1B) (*12, 13*). The homology spans both direct repeats that form a bipartite molecular saddle structure in eukaryotic TBPs. Notably, strong conservation exists at residues that are critical for DNA binding in eukaryotic TBPs. Of the 35 residues that mediate DNA contacts in the TBP crystal structures of *Saccharomyces cerevisiae* or of *Arabidopsis thaliana* (or both) (*13*), 25 are identical in *P. woesei* TBP, and the rest represent conservative changes (Fig. 1B).

These observations suggest that *P. woesei* TBP binds DNA in a manner similar to that of eukaryotic TBPs. Indeed, using the electrophoretic mobility-shift assay (EMSA), we found that purified recombinant *P. woesei* TBP bound specifically to an oligonucleotide containing the TATA box–like A-box region of a *P. woesei* gene (*14*) (Fig. 2B). The DNA binding was essentially unaffected by treatment of the protein for 30 min at 70°C. This is consistent with the fact that *P. woesei* is a hyperthermophile that thrives at over 100°C and contrasts with eukaryotic TBP proteins, which are inactivated by heating to 47°C (*15*). Thus, the *P. woesei* gene we have isolated encodes a bona fide TATA-binding protein.

In eukaryotes, complexes between TBP and DNA are stabilized by the subsequent recruitment of general transcription factor TFIIB (*16, 17*). Previously, database searches have led to the identification of a putative *P. woesei* TFIIB homolog (*8*). We therefore wished to determine whether this factor could interact with *P. woesei* TBP in the EMSA. Because archaebacteria contain only one RNA polymerase and because the *P. woesei* TFIIB homolog is, in fact, equally related to the RNA polymerase III general factor BRF1 (*8, 17, 18*), we refer to it as transcription factor B (TFB). Although *P. woesei* TFB does not bind DNA on its own in detectable amounts, when it was incubated with *P. woesei* TBP and DNA, a distinct TBP-TFB-DNA complex was generated (Fig. 2C). The formation of this complex was stimulated by elevated temperatures—the autoradiographic exposure was much shorter at 55°C than at 30°C (Fig. 2C). Under these conditions, *P. woesei* TBF must increase either the rate of assembly or the stability of TBP-DNA complexes. The *P. woesei* TBP-TFB-DNA complex was subject to competition by an oligonucleotide bearing the TATA-like A-box *Pyrococcus* gene promoter region but not by an oligonucleotide derivative bearing mutations in this sequence (Fig. 2D). In terms of forming complexes on promoter DNA, *Pyrococcus* TBP and TFB thus behave in an analogous fashion to eukaryotic TBP and TFIIB, respectively, which is consistent with the notion that these archaebacterial and eukaryotic factors are functionally homologous. However, *P. woesei* TBP is unable to substitute for human TBP in transcription reactions in vitro (*19*), presumably as a result of the huge evolutionary gulf between *Pyrococcus* and humans and because of the different growth optima for these two organisms.

In eukaryotes, TBP serves as a target for transcriptional regulators (*2, 3*). In some cases, transcriptional activation or repression by DNA-bound regulators appears to

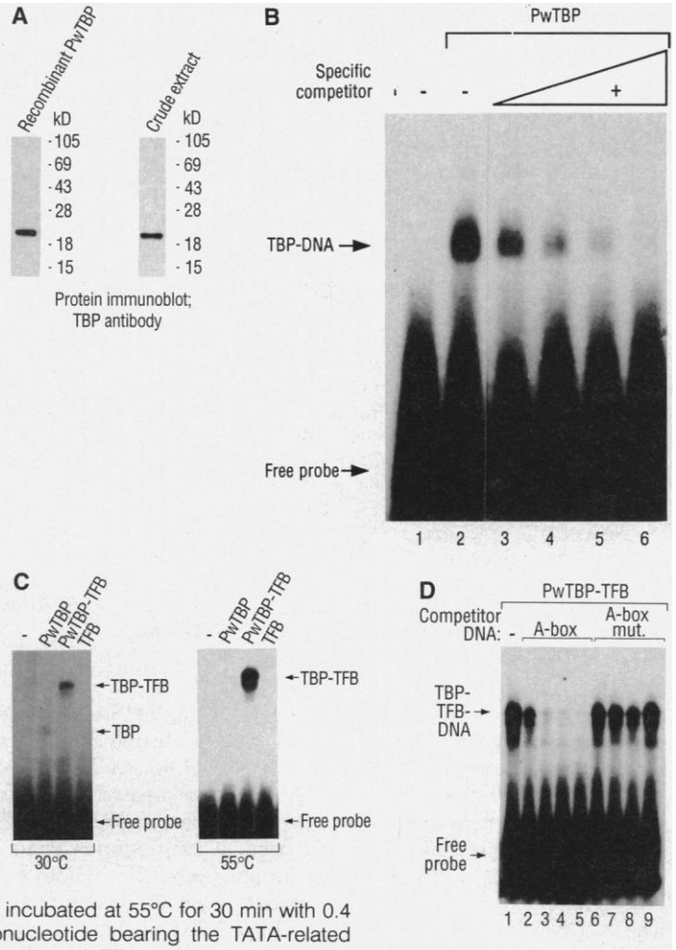

Fig. 2. (**A**) Detection of *P. woesei* TBP in *Pyrococcus* cell extracts. The gene encoding PwTBP was cloned into *E. coli* expression vector pQE30 (Qiagen, Chatsworth, California) and purified to over 98% homogeneity by Ni-nitrolotriacetic acid–affinity chromatography (*21*). The PwTBP protein (200 ng) and 150 μg of crude *Pyrococcus* cell extract were electrophoresed on a 12% SDS-polyacrylamide gel, blotted onto nitrocellulose, and then probed with TBP monoclonal antibody 16E8. The recombinant PwTBP was synthesized fused to 18 extra amino acid residues, including six histidines, and therefore migrated slightly slower than the endogenous protein. (**B**) Specific DNA binding by PwTBP. The EMSA conditions were those optimized for eukaryotic TBPs (*16, 17*). Lane 1: Probe alone. Lanes 2 to 6: Purified recombinant PwTBP was incubated at 55°C for 30 min with 0.4 ng of radiolabeled oligonucleotide bearing the TATA-related A-box region of the *P. woesei* EF1α gene promoter (*14*) (upper strand sequence, 5′-GCGAAAGCTTTAAAAAGTAAGTTCAAA-3′) together with 50 ng of nonspecific competitor poly-(dG):poly(dC). Incubations were either in the absence (lane 2) or presence of unlabeled oligonucleotides (2.5-, 12-, 25-, and 60-fold excess over radiolabeled probe; lanes 3 to 6, respectively). Parallel experiments with unrelated oligonucleotides or the A-box–mutant oligonucleotide [see (C) and (D)] showed no competition. (**C**) PwTBP and TFB bound DNA cooperatively and showed increased activity at elevated temperatures. Purified recombinant PwTBP or TFB (or both) were incubated with radiolabeled A-box oligonucleotide probe at the indicated temperature and analyzed as in (B). Dashed lane indicates no protein added. (**D**) The PwTBP-TFB complex bound specifically to the TATA-related A-box sequence found in *Pyrococcus* promoters. The PwTBP-TFB complexes were formed at 55°C as in (C). Reactions contained as competitor either the A-box oligonucleotide (lanes 2 to 5) or an A-box–mutant oligonucleotide (lanes 6 to 9; sequence 5′-GCGAAAGCTTTCCCAAGTAAGTTCAAA-3′). Competitor: probe ratios were as in (B). Dashed lane (lane 1) indicates no competitor DNA added.

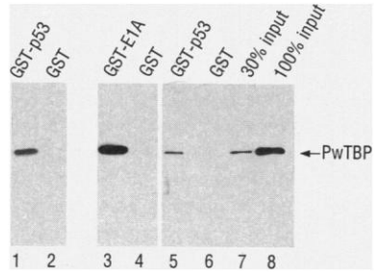

Fig. 3. PwTBP interacts with eukaryotic transcriptional activation domains. Generation and purification of glutathione-S-transferase (GST)–fusion proteins was as described (*22*); equal quantities of each protein were used. Samples of glutathione-Sepharose beads bearing GST alone, GST-E1A (full-length 13*S* E1A), or GST-p53 (residues 1 to 73) were preincubated at room temperature for 15 min with recombinant full-length PwTBP (lanes 1 and 2) or with PwTBP residues 52 to 191 (lanes 3 to 8) in the presence of 30 μg of bovine serum albumin. After extensive washing, beads were mixed with SDS-sample buffer and samples electrophoresed on a 12% SDS-polyacrylamide gel. Bound TBP was detected by protein immunoblot with the use of TBP monoclonal antibody 16E8.

be mediated by them contacting TBP directly. A region of TBP that is implicated in these interactions encompasses the helix-2 and basic repeat structures at the end of the first direct repeat (*3, 12, 13*). Because *P. woesei* TBP is predicted to possess an analogous region, we tested whether it could interact with eukaryotic activation domains known to bind eukaryotic TBPs. *Pyrococcus woesei* TBP bound tightly and specifically to activation domains derived from the adenovirus E1A protein and from the anti-oncogene protein p53 (Fig. 3). These findings suggest that eukaryotic and *P. woesei* TBPs are analogous functionally and suggest that eukaryotic-type transcriptional activators are present in archaebacteria.

Taken together with previous studies, our findings suggest that the transcriptional machineries of archaebacteria and eukaryotes are fundamentally homologous; they are also consistent with phylogenetic comparisons that place eukaryotes closer to archaebacteria than to eubacteria (*6, 7*). The eukaryotic-type transcriptional apparatus must therefore have already been established in the last common ancestor of eukaryotes and archaebacteria, before the emergence of nucleated cells. These considerations lead to the prediction that other eukaryotic transcriptional components (such as TAFs, TFIIA, TFIIE, and TFIIF) will also exist in archaebacteria. Studying archaebacterial transcription will be valuable in increasing our understanding of these organisms and promises to provide insights into the evolution and functioning of the eukaryotic nuclear transcriptional apparatus.

REFERENCES AND NOTES

1. R. J. White and S. P. Jackson, *Trends Genet.* **8**, 284 (1992); N. Hernandez, *Genes Dev.* **7**, 1291 (1993); P. W. J. Rigby, *Cell* **72**, 7 (1993).
2. J. A. Inostroza, G. H. Mermelstein, I. Ha, W. S. Lane, D. Reinberg, *Cell* **70**, 477 (1992).
3. X. Liu, C. W. Miller, P. H. Koeffler, A. J. Berk, *Mol. Cell. Biol.* **13**, 3291 (1993); C. J. Ingles, M. Shales, W. D. Cress, S. J. Triezenberg, J. Greenblatt, *Nature* **351**, 588 (1991); E. Seto *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 12028 (1992); N. Horikoshi *et al.*, *ibid.* **88**, 5124 (1991); W. S. Lee, C. C. Kao, G. O. Bryant, X. Liu, A. J. Berk, *Cell* **67**, 365 (1991).
4. J. D. Fikes, D. M. Becker, F. Winston, L. Guarente, *Nature* **346**, 291 (1990); A. Hoffmann *et al.*, *ibid.*, p. 387; A. Hoffmann *et al.*, *Genes Dev.* **4**, 1141 (1990); C. C. Kao *et al.*, *Science* **248**, 1646 (1990); T. Hoey, B. D. Dynlacht, M. G. Peterson, B. F. Pugh, R. Tjian, *Cell* **61**, 1179 (1990); T. Yamamoto *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2844 (1992); M. B. McAndrew, M. Read, P. F. G. Sims, J. E. Hyde, *Gene* **124**, 165 (1993).
5. L. A. Allison, M. Moyle, M. Shales, C. J. Ingles, *Cell* **42**, 599 (1985); C. Carles, I. Treich, F. Bouet, M. Riva, A. Sentenac, *J. Biol. Chem.* **266**, 24092 (1991).
6. C. R. Woese, O. Kandler, M. L. Wheelis, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4576 (1990); P. Cammarano *et al.*, *J. Mol. Evol.* **34**, 396 (1991); S. P. Jackson, *Trends Genet.*, in press.
7. G. Pühler *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4569 (1989); H.-P. Klenk, P. Palm, F. Lottspeich, W. Zillig, *ibid.* **89**, 407 (1992).
8. C. Ouzounis and C. Sander, *Cell* **71**, 189 (1992); R. Creti, P. Londei, P. Cammarano, *Nucleic Acids Res.* **21**, 2942 (1993).
9. C. D. Langer and W. Zillig, *Nucleic Acids Res.* **21**, 2251 (1993).

10. M. Thomm and G. Wich, *ibid.* **16**, 151 (1988); W.-D. Reiter, U. Hüdepohl, W. Zillig, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9509 (1990); W. Hausner, G. Frey, M. Thomm, *J. Mol. Biol.* **222**, 495 (1991); J. Hain, W.-D. Reiter, U. Hüdepohl, W. Zillig, *Nucleic Acids Res.* **20**, 5423 (1992); U. Hüdepohl, W.-D. Reiter, W. Zillig, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5851 (1990).
11. R. J. White and S. P. Jackson, *Cell* **71**, 1041 (1992).
12. D. B. Nikolov *et al.*, *Nature* **360**, 40 (1992); D. I. Chasman, K. M. Flaherty, P. A. Sharp, R. D. Kornberg, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8174 (1993).
13. Y. Kim, J. H. Geiger, S. Hahn, P. B. Sigler, *Nature* **365**, 512 (1993); J. L. Kim, D. B. Nikolov, S. K. Burley, *ibid.*, p. 520; S. Burley, personal communication.
14. R. Creti *et al.*, *J. Mol. Evol.* **33**, 332 (1991).
15. N. Nakajima, M. Horikoshi, R. G. Roeder, *Mol. Cell. Biol.* **8**, 4028 (1988); R. J. White, S. P. Jackson, P. W. J. Rigby, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1949 (1992).
16. S. Buratowski, S. Hahn, L. Guarente, P. A. Sharp, *Cell* **56**, 549 (1989); E. Maldonado, I. Ha, P. Cortes, L. Weis, D. Reinberg, *Mol. Cell. Biol.* **10**, 6335 (1990); K. Hisatake, R. G. Roeder, M. Horikoshi, *Nature* **363**, 744 (1993).
17. S. Buratowski and H. Zhou, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5633 (1993).
18. ———, *Cell* **71**, 221 (1992); T. Colbert and S. Hahn, *Genes Dev.* **6**, 1940 (1992); A. Lopez-de-Leon, M. Librizzi, K. Tuglia, I. Willis, *Cell* **71**, 211 (1992).
19. R. J. White, P. Baumann, S. P. Jackson, unpublished results.
20. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
21. R. Janknecht *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8972 (1991).
22. A. J. Bannister, T. M. Gottlieb, T. Kouzarides, S. P. Jackson, *Nucleic Acids Res.* **21**, 1289 (1993).
23. We are indebted to P. Cammarano for providing archaebacterial genomic DNAs; to N. Finnie, A. Bannister, and T. Kouzarides for the GST protein derivatives; to R. Weinzierl and R. Tjian for the TBP monoclonal antibody 16E8; and to G. Ghosh and P. Sigler for the *P. woesei* TFB expression plasmid. We also thank M. McAndrew and J. Hyde for advice on PCR conditions, R. Dendy for secretarial assistance, and members of S.P.J.'s laboratory for advice and encouragement. The laboratory of S.P.J. is funded principally by project grant SP2143/0101 from the Cancer Research Campaign (UK).

19 October 1993; accepted 29 April 1994

Candidate Gene Associated with a Mutation Causing Recessive Polycystic Kidney Disease in Mice

Judith H. Moyer, Monica J. Lee-Tischler, Heajoon-Y. Kwon, Jeffrey J. Schrick, Ellis D. Avner, William E. Sweeney, Virginia L. Godfrey, Nestor L. A. Cacheiro, J. Erby Wilkinson, Richard P. Woychik*

A line of transgenic mice was generated that contains an insertional mutation causing a phenotype similar to human autosomal recessive polycystic kidney disease. Homozygotes displayed a complex phenotype that included bilateral polycystic kidneys and an unusual liver lesion. The mutant locus was cloned and characterized through use of the transgene as a molecular marker. Additionally, a candidate polycystic kidney disease (PKD) gene was identified whose structure and expression are directly associated with the mutant locus. A complementary DNA derived from this gene predicted a peptide containing a motif that was originally identified in several genes involved in cell cycle control.

Autosomal recessive PKD (ARPKD) typically affects infants and children at an estimated incidence of up to 1 in 10,000 live births (*1*), although prenatal deaths and stillbirths caused by ARPKD may increase that incidence by 27 times (*2*). One notable difference between recessive and dominant PKD (ADPKD), which affects up to 1 in 500 individuals (*1*), is the nearly invariable presence in the former of noncystic lesions of the liver, ranging from biliary dysplasia to portal fibrosis (*3*).

Efforts to identify the human chromosomal regions associated with PKD are currently underway. The locus responsible for more than 85% of ADPKD (*PKD1*) has been localized to human 16p13.3 (*4*), and non-*PKD1* forms of ADPKD have been linked to additional chromosomes (*5*). The entire *PKD1* locus has been isolated on a series of overlapping clones (*6*), and although at least 23 genes have been identified in this 750-kb region (*7*), it still remains unclear which of these genes, if any, is associated with ADPKD.

Animal models of PKD that resembles the various forms of human PKD have been extremely useful for studying the pathogenesis of cystic kidney disease. The recessive *cpk* (congenital polycystic kidneys) mutation (*8*) has been examined in considerable detail (*9*, *10*). More recently, additional spontaneous recessive mutations in mice—*bpk* (BALB/c polycystic kidneys) (*11*), *pcy* (polycystic kidneys) (*12*), and *jck* (juvenile cystic kidneys) (*13*)—have been reported. The cystic disease that develops in the *cpk*, *bpk*, and *jck* models shows similarities to clinical features of human ARPKD, but the dual hepatorenal pathology that is a hallmark of ARPKD has been reported as a consistent feature only in the *bpk* mouse (*11*). The *cpk* locus has been mapped to a region of mouse chromosome 12 with homology to human chromosome 2p24-2p25 (*14*), and the *pcy* locus to a region of mouse chromosome 9 homologous to human chromosome 3 (*12*). In none of these models has the mutant gene been identified, primarily because of a lack of DNA probes directly associated with the locus. Therefore, at present it is impossible to test directly whether the gene associated with any of these mutations corresponds to a gene in humans that causes ARPKD.

Here we report the molecular, genetic, and phenotypic characterization of an insertional mutation in mice that gives rise to a recessive form of PKD that has marked similarities to human ARPKD. Unlike other mutations giving rise to cystic kidneys in mice, this mutation arose by insertional mutagenesis and is called *TgN(Imorpk)-737Rpw* (*Imorpk*: insertional mutation, Oak Ridge polycystic kidneys) [named in accordance with the new nomenclatural rules for transgenic mice as described (*15*), and abbreviated here as *TgN737Rpw*]. The *TgN737Rpw* mutation arose as part of a large-scale insertional mutagenesis program at the Oak Ridge National Laboratory. The *TgN737Rpw* transgenic line was generated on the FVB/N inbred genetic background by pronuclear microinjection (*16*) of the linearized pPyF9-1 construct, which contains the bacterial chloramphenicol acetyltransferase gene under the control of a mutated version of the polyoma early region promoter (*17*).

The abnormal phenotype arose in the *TgN737Rpw* line only among transgenic homozygotes, derived from crosses between animals heterozygous for the transgene. In the FVB/N inbred genetic background, external examination revealed that mutant animals have scruffy fur, are severely growth retarded, and have preaxial polydactyly on all limbs. In an analysis of 159 live offspring from a cross of (C3H × FVB/N)-*TgN737Rpw* heterozygotes, mice from 33 complete litters were genotyped for the presence of the transgene. Thirty-five

J. H. Moyer, M. J. Lee-Tischler, H.-Y. Kwon, J. J. Schrick, University of Tennessee Graduate School of Biomedical Sciences, Biology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831–8077, USA.
E. D. Avner and W. E. Sweeney, Department of Pediatrics, University of Washington, Seattle, WA 98105, USA.
V. L. Godfrey, N. L. A. Cacheiro, R. P. Woychik, Biology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831–8077, USA.
J. E. Wilkinson, Department of Pathobiology, University of Tennessee, Knoxville, TN 37920, USA.
*To whom correspondence should be addressed.