Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 4279–4283, May 1995
Biochemistry

Replication factor encoded by a putative oncogene, *set*, associated with myeloid leukemogenesis

(adenovirus/leukemia/cDNA cloning/nucleosome assembly protein)

KYOSUKE NAGATA*†, HIROYUKI KAWASE*, HIROSHI HANDA*, KEIICHI YANO‡, MOTOO YAMASAKI‡, YUKIO ISHIMI§, ASUKO OKUDA§, AKIHIKO KIKUCHI§, AND KEN MATSUMOTO§

*Department of Biomolecular Engineering, Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama 226, Japan; ‡Tokyo Research Laboratories, Kyowa Hakko, Machida, Tokyo 194, Japan; and §Mitsubishi Kasei Institute of Life Sciences, Machida, Tokyo 194, Japan

Communicated by Jerard Hurwitz, Memorial Sloan–Kettering Cancer Center, New York, NY, January 27, 1995

ABSTRACT **DNA replication of the adenovirus genome complexed with viral core proteins is dependent on the host factor designated template activating factor I (TAF-I) in addition to factors required for replication of the naked genome. Recently, we have purified TAF-I as 39- and 41-kDa polypeptides from HeLa cells. Here we describe the cloning of two human cDNAs encoding TAF-I. Nucleotide sequence analysis revealed that the 39-kDa polypeptide corresponds to the protein encoded by the *set* gene, which is the part of the putative oncogene associated with acute undifferentiated leukemia when translocated to the *can* gene. The 41-kDa protein contains the same amino acid sequence as the 39-kDa protein except that short N-terminal regions differ in both proteins. Recombinant proteins, which were purified from extracts of *Escherichia coli*, expressing the proteins from cloned cDNAs, possessed TAF-I activities in the *in vitro* replication assay. A particular feature of TAF-I proteins is the presence of a long acidic tail in the C-terminal region, which is thought to be an essential part of the SET–CAN fusion protein. Studies with mutant TAF-I proteins devoid of this acidic region indicated that the acidic region is essential for TAF-I activity.**

DNA replication and transcription of adenovirus (Ad) have been studied extensively as a model eukaryotic system. The dissection and reconstitution of the cell-free DNA replication system using the Ad DNA terminal protein complex (Ad DNA-prot) have revealed the detailed mechanism of Ad genome replication (1–3). The Ad genome is a linear DNA of ≈36 kbp that contains 55-kDa terminal proteins covalently attached to its 5′ ends. Replication of the Ad DNA-prot initiates by a protein-priming mechanism in which the 5′ terminal nucleotide of the nascent DNA, dCMP, is linked to the 80-kDa precursor of the terminal protein (pTP). The nontemplate strand is displaced as the nascent DNA chain elongates. Three virus-encoded proteins—the Ad DNA polymerase, the Ad DNA binding protein, and pTP—are essential for initiation of Ad DNA-prot replication and the elongation reaction. NFI, a host factor, stimulates formation of the pTP–dCMP initiation complex as well as elongation of DNA up to 30% of the length of full-sized Ad DNA (4). Another host factor, NFIII, also stimulates initiation (5). NFI and NFIII have been shown to be sequence-specific DNA binding proteins that bind to their cognate binding sites within the inverted terminal repeat of Ad DNA (5, 6). NFII, a type I topoisomerase, is required for synthesis of full-length Ad DNA in the presence of three viral proteins, NFI, and NFIII (20).

The Ad genome in virions and in infected cells at early stages of infection exists in the form of the complex of Ad DNA and viral basic core proteins (7, 8). This DNA–protein complex, called Ad core, is thought to be a bona fide template for transcription of early genes and genome replication. Since DNA replication from the Ad core does not occur in the *in vitro* system with purified proteins required for replication of the naked Ad genome (9, 10), we have developed an *in vitro* system in which the Ad core can function as a template for replication (10). With this system, we found a stimulatory activity for the Ad core DNA replication in uninfected HeLa cell extracts. The factor stimulating Ad core DNA replication, designated template activating factor I (TAF-I), was purified as a protein with a molecular mass of either 41 kDa (TAF-Iα) or 39 kDa (TAF-Iβ) (10).

In an attempt to define the function of TAF-I, here we describe cloning of the cDNAs encoding TAF-I. Nucleotide sequence analysis has revealed that TAF-I polypeptides are encoded by the putative oncogene *set* (11). Studies with recombinant TAF-I proteins indicate that the carboxyl acidic region of the TAF-I protein is required for stimulation of Ad genome replication. In addition, we show that the nucleosome assembly protein (NAP), which has significant amino acid sequence similarity to TAF-I, substitutes for TAF-I activity in the cell-free Ad core replication system.¶

MATERIALS AND METHODS

Cell-Free DNA Replication. The Ad DNA replication assay and purification of Ad core and TAF-I were performed as described (10). Mouse NAP-I was purified from extracts of *Escherichia coli* expressing the protein from its cDNA (isolated by A.O. and A.K.) as described for purification of yeast NAP-I (12).

Cloning of TAF-I cDNAs. Each TAF-Iα or -Iβ protein (≈10 pmol) was separated by PAGE in the presence of SDS, eluted from the gel, and subjected to digestion with lysylendopeptidase. Digested peptides were separated by reverse-phase HPLC. Amino acid sequences were determined with the gas-phase sequencer (Applied Biosystems, model 470A/120A) using one peptide unique to TAF-Iα and four peptides common to TAF-Iα and -Iβ. Degenerate oligonucleotides, corresponding to the N-terminal and C-terminal parts of the peptide shown by bracket 2 of Fig. 1, were synthesized and used as primers for reverse transcriptase-mediated PCR with HeLa cell mRNA. Nucleotide sequence analysis of 60-bp amplified DNA cloned into vector plasmid revealed that the cloned DNA encoded the amino acid sequence of peptide 2. Using this DNA fragment, a λ HeLa cDNA library obtained from Clontech was screened. Positive clones were plaque-purified twice and insert DNAs were subcloned into plasmids and sequenced.

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: TAF-I, template activating factor I; Ad, adenovirus; NAP, nucleosome assembly protein; r, recombinant.

†To whom reprint requests should be addressed.

¶The sequences reported in this paper have been deposited in the GenBank data base (accession nos. D12518 and D45198).

Preparation of Recombinant (r) TAF-I Proteins. To obtain a large amount of rTAF-I, a T7 expression system of *E. coli* was used. For generation of inserts for pET14b, cloned TAF-I cDNAs were used for amplification by PCR with oligonucleotide primers as follows: either 5′-GGCAGCCATATGGCCCCTAAACGCCAGTCTCCA-3′ as an N-terminal primer for TAF-Iα or 5′-GGCAGCCATATGTCGGCGCCGGCGGCCAAAGTC-3′ as an N-terminal primer for TAF-Iβ and 5′-CGCGGATCCTTAGTCATCTTCTCCTTCATC-3′ for both rTAF-Iα(1–290) and rTAF-Iβ(1–277), 5′-CGCGGATCCTTATCCTTCCTCCCCTTCATC-3′ for both rTAF-Iα(1–282) and rTAF-Iβ(1–269), 5′-CGCGGATCCTTAAATATCTTCTAATCCTTC-3′ for both rTAF-Iα(1–261) and rTAF-Iβ(1–248), or 5′-CGCGGATCCTTACATATCGGGAACCAAGTA-3′ for both rTAF-Iα(1–238) and rTAF-Iβ(1–225) as C-terminal primers were used. Amplified DNAs were digested with both *Nde* I and *Bam*HI and cloned into *Nde* I and *Bam*HI-digested pET14b. BL21(DE3) was transformed with each plasmid. A 10-ml culture in L-broth was induced by the addition of isopropyl β-D-thiogalactopyranoside to synthesize rTAF-I proteins. Sonicated cell lysates were subjected to the purification system using Ni-chelation resins according to the method suggested by the manufacturer (Novagen). Approximately 80 μg of rTAF-I was prepared. Further isolation of each peptide was carried out from gel slices containing TAF-I purified from HeLa cells or rTAF-I proteins by the denature–renature protocol described (13).

RESULTS AND DISCUSSION

Cloning of cDNA for TAF-I. To aid in cloning the cDNAs encoding TAF-Iα and -Iβ proteins, each protein was isolated and digested with lysylendopeptidase. Digested peptides were separated and subjected to amino acid sequence analysis. Four amino acid sequences common to both proteins and one sequence unique to TAF-Iα were determined (Fig. 1*B*). Based on one of the common amino acid sequences, two degenerate oligonucleotides were synthesized and used for reverse transcriptase-mediated PCR with HeLa cell mRNA. Cloning and sequencing of amplified DNA revealed that the cloned DNA encoded the determined amino acid sequence. We used this cloned DNA as a probe for screening a HeLa cDNA library. Two different cDNA clones were isolated after screening ≈500,000 recombinants. Sequencing of the clones showed that one, encoding four common peptides, contained the same nucleotide sequence as the *set* gene (11) and the other contained the nucleotide sequence (Fig. 1*A*) that encoded four common peptides and one peptide unique to TAF-Iα. These findings suggested that the polypeptides encoded by the former and the latter cDNA clones were TAF-Iβ (or SET) and TAF-Iα, respectively (Fig. 1*B*). Although TAF-Iβ was thought to be generated from TAF-Iα by proteolytic cleavage, it is possible that both proteins are synthesized from alternatively spliced mRNAs. The encoded polypeptides have predicted molecular masses of 34 and 32 kDa for TAF-Iα and -Iβ, respectively. The difference in molecular masses between the observed and predicted proteins is presently unclear, but it may be due to the fact that the *set* gene product is phosphorylated (14) and/or that these proteins have long acidic regions (see below), which contributes to unusual migration on PAGE in the presence of SDS (14).

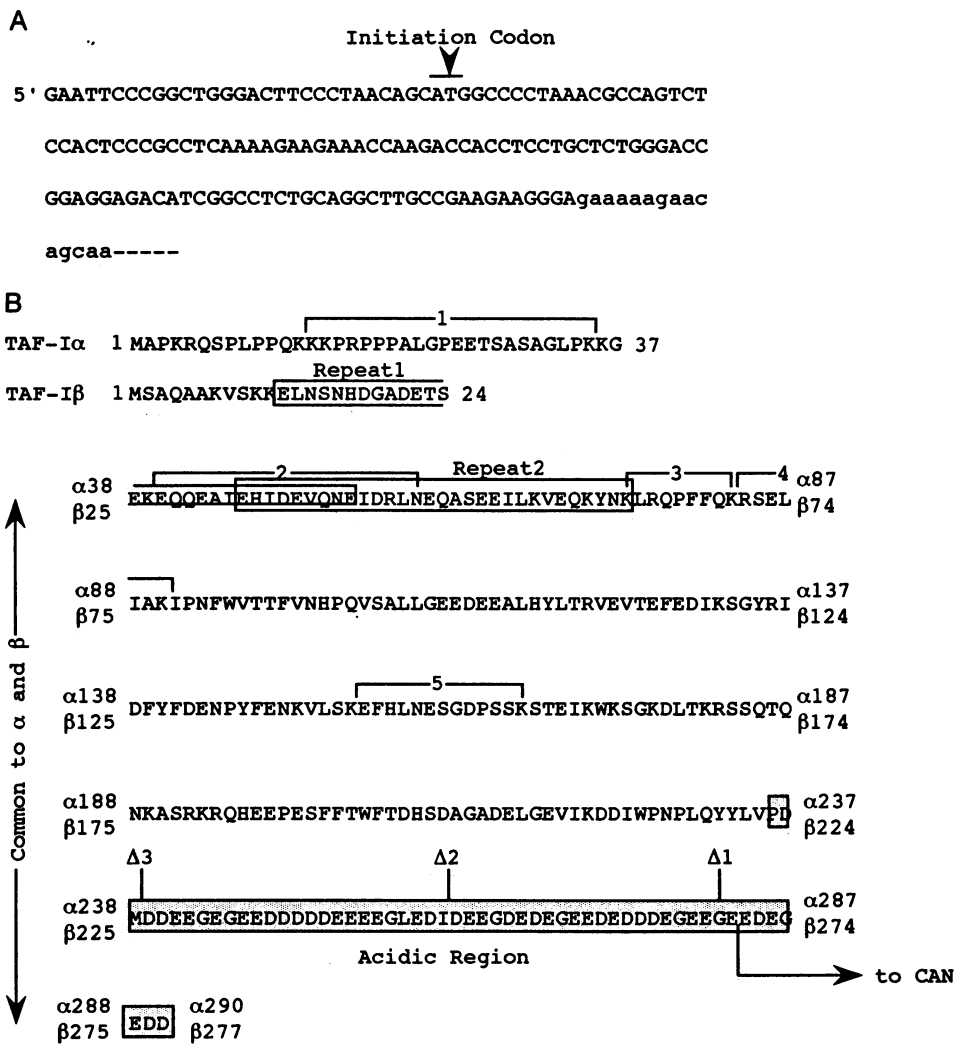

FIG. 1. (*A*) Nucleotide sequence of human TAF-Iα cDNA. A portion of the 5′ terminal sequence of cloned TAF-Iα cDNA is shown. Nucleotide sequences indicated by uppercase and lowercase letters are the sequence specific for TAF-Iα and that common to both TAF-Iα and -Iβ, respectively. A putative initiation ATG codon is overlined. (*B*) Deduced amino acid sequences of human TAF-Iα and -Iβ. Numbering is from putative initiation methionines of TAF-Iα and -Iβ deduced from their respective cDNAs (*A*; see ref. 11). Amino acid sequences determined from peptides isolated from purified TAF-Iα and -Iβ are indicated by brackets 1–5. The intramolecular overlapping repeat regions containing a low but distinct similarity in amino acid sequences are shown by Repeat 1 and Repeat 2. The long acidic tail is highlighted by the shaded box. Bars with Δ1–3 indicate C termini of mutant TAF-I proteins used in Fig. 2. The translocation breakpoint fused to the CAN found in acute undifferentiated leukemia (11) is indicated by an arrow.

The *set*/TAF-Iβ gene was identified as a fusion gene with the *can* gene in a translocated region found in acute undifferentiated leukemia (11). CAN contains putative dimerization motifs and a possible DNA binding domain (15). The *can* gene is also found in the form of the *dek–can* fusion gene in acute myeloid leukemia (14). There is no significant amino acid sequence homology between SET and DEK except that both contain highly acidic regions. In fact, CAN fused with SET or DEK, both of which contain the acidic regions, functions as an oncogene. The fusion point of SET/TAF-Iβ to CAN and the long acidic region of TAF-I, which were previously noted (11, 15), are indicated in Fig. 1*B*. By additional computer-assisted analysis, we detected overlapping 30-amino acid repeats in TAF-Iβ (Repeat 1 and Repeat 2 in Fig. 1*B*).

Determination of Functional Domain for TAF-I Activity. To determine the functional domain of TAF-I in the Ad core DNA replication reaction, we prepared mutant TAF-I proteins. In particular, mutants devoid of the acidic region were constructed and their ability to support *in vitro* DNA replication was examined. Using the T7 expression system in *E. coli*, histidine-tagged wild-type TAF-Iα and -Iβ (Fig. 2*B*), designated rTAF-Iα(1–290) and rTAF-Iβ(1–277), respectively (numbers in parentheses indicate amino acid residues), were prepared. Both recombinant TAF-I proteins showed stimulatory activity in the replication assay (Fig. 2*A*), indicating that the cloned cDNAs encoded TAF-I. The specific TAF-I activity of rTAF-Iβ is higher than that of rTAF-Iα (Fig. 2*A*, lanes 9–14), being in good agreement with the results observed when

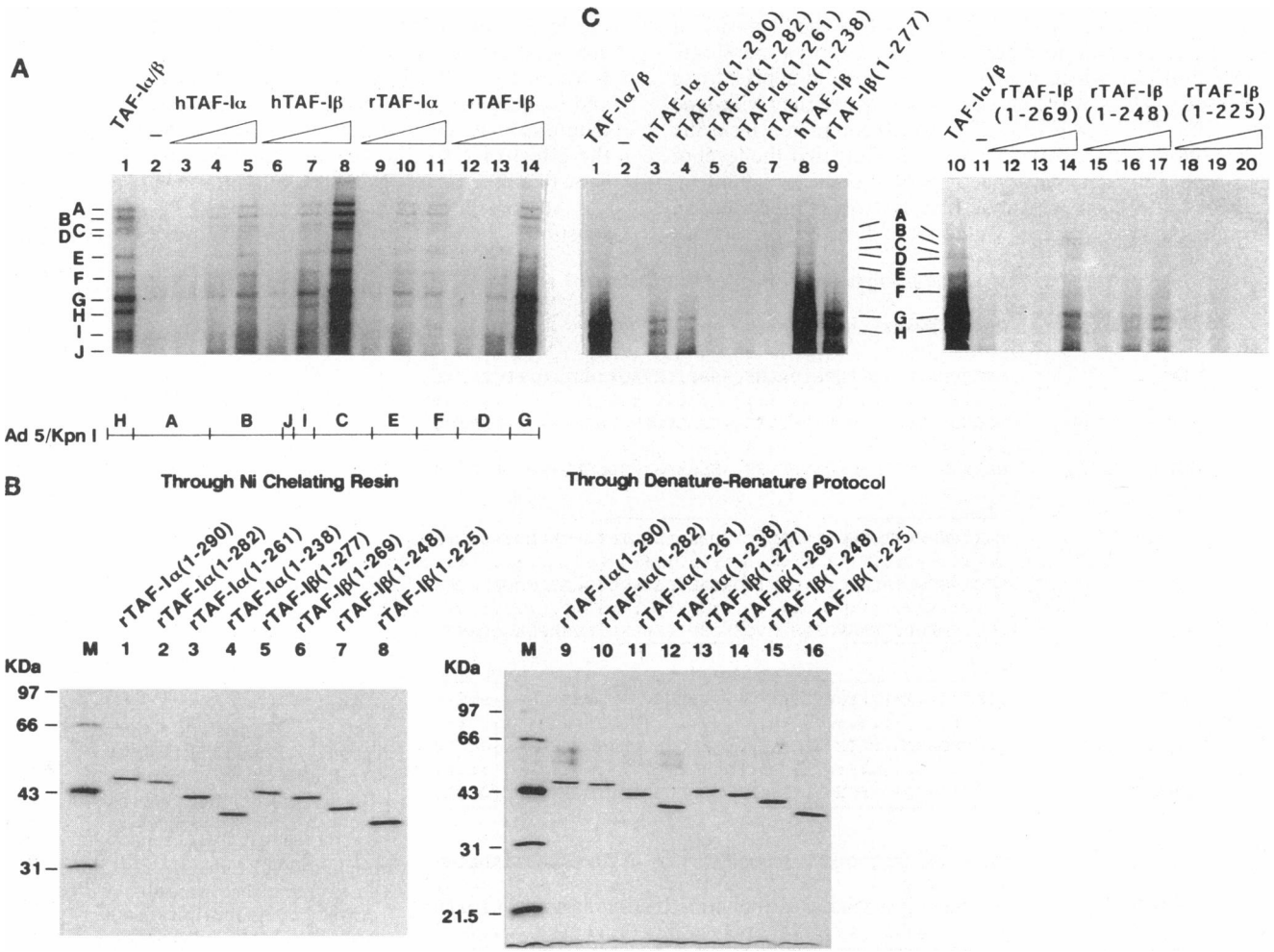

FIG. 2. (*A*) Stimulatory activity for the Ad core DNA replication of recombinant TAF-I. The Ad core DNA replication assay was carried out as described (10) with factors required for replication of naked Ad genome in the absence (lane 2) or presence of 30 ng of native TAF-I purified from HeLa cells (lane 1) as a positive control, or 2 ng (lanes 3, 6, 9, and 12), 10 ng (lanes 4, 7, 10, and 13), or 50 ng (lanes 5, 8, 11, and 14) of isolated human (h)TAF-Iα (lanes 3–5), hTAF-Iβ (lanes 6–8), rTAF-Iα(1–290) (lanes 9–11), or rTAF-Iβ(1–277) (lanes 12–14), which were purified from gel slices and subjected to the denature–renature protocol (13). For analysis of the products by gel electrophoresis, reactions were terminated by SDS addition and proteinase K digestion. DNA products were purified, digested with *Kpn* I, and separated by electrophoresis through a 0.8% agarose gel. The gel was dried and autoradiographed. *Kpn* I fragments of Ad type 5 DNA are indicated on the left of the gel and positions of *Kpn* I fragments in the Ad genome are indicated at the bottom. *Kpn* I G and H fragments are located at the ends of the Ad genome DNA where Ad DNA replication starts. (*B*) Purification of recombinant TAF-I proteins. rTAF-I proteins [rTAF-Iα(1–290), rTAF-Iα(1–282), rTAF-Iα(1–261), and rTAF-Iα(1–238) and rTAF-Iβ(1–277), rTAF-Iβ(1–269), rTAF-Iβ(1–248), and rTAF-Iβ(1–225)], tagged with histidine, were purified from *E. coli* by using Ni-chelate resins (lanes 1–8). rTAF-I proteins were further purified from gel slices by a denature–renature protocol (13) (lanes 9–16). Proteins were subjected to PAGE in the presence of SDS and visualized by silver staining. Molecular weight markers (lanes M) were obtained from Bio-Rad. (*C*) Stimulatory activity of rTAF-I proteins for Ad core DNA replication. The Ad core DNA replication assay was performed as described above in the absence (lanes 2 and 11) or presence of 20 ng of native TAF-I purified from HeLa cells (lanes 1 and 10), 20 ng of isolated hTAF-Iα (lane 3), 20 ng of hTAF-Iβ (lane 8), 20 ng of rTAF-Iα(1–290) (lane 4), 20 ng of rTAF-Iα(1–282) (lane 5), 20 ng of rTAF-Iα(1–261) (lane 6), 20 ng of rTAF-Iα(1–238) (lane 7), 20 ng of rTAF-Iβ(1–277) (lane 9), or 2 ng (lanes 12, 15, and 18), 10 ng (lanes 13, 16, and 19), or 50 ng (lanes 14, 17, and 20) of rTAF-Iβ(1–269) (lanes 12–14), rTAF-Iβ(1–248) (lanes 15–17), or rTAF-Iβ(1–225) (lanes 18–20). All hTAF-Iα and -Iβ and rTAF-Iα and -Iβ derivatives were purified by the denature–renature protocol (13).

isolated native TAF-I proteins were used (Fig. 2*A*, lanes 3–8). These results suggest that the N-terminal region, specific for each protein, influences the specific activity of these proteins. A mixing experiment with TAF-Iα and -Iβ proteins resulted in an additive effect on DNA replication assay (data not shown). Next, we examined the effect of amino acid deletion from rTAF-I on its stimulatory activity in the Ad core DNA replication system (Fig. 2*C*). Three different mutant proteins of TAF-Iα or TAF-Iβ were constructed, purified (Fig. 2*B*), and used in the DNA replication assay—i.e., rTAF-Iα(1–282), rTAF-Iα(1–261), and rTAF-Iα(1–238) or rTAF-Iβ(1–269), rTAF-Iβ(1–248), and rTAF-Iβ(1–225). rTAF-Iα(1–282) retained only a trace amount of TAF-I activity compared with that of wild-type rTAF-Iα(1–290). rTAF-Iα(1–261) and rTAF-Iα(1–238) were virtually inactive. rTAF-Iβ(1–269) and rTAF-Iβ(1–248) showed 59% and 43% of the TAF-I activity relative to that of wild-type rTAF-Iβ(1–277), when the intensity of bands corresponding to fragments G and H was measured. rTAF-Iβ(1–269), which is only one amino acid shorter than SET fused to CAN, was considerably less active than wild-type rTAF-Iβ(1–277). rTAF-Iβ(1–225), which completely lacks the acidic region, was inactive (<5%). It is likely that the level of DNA replication stimulatory activity of TAF-Iα, which intrinsically possesses a low specific activity, is strongly dependent on the complete acidic region, while that of TAF-Iβ corresponds to the length of the acidic region. These results indicate that the functional domain of TAF-I required for the DNA replication stimulatory activity resides in the acidic region.

Structural and Functional Relationship Between TAF-I and Other Proteins. No statistically significant similarity to entries in the Protein Identification Resource data bank was found for TAF-I. However, TAF-I shows a low but distinct sequence similarity with NAPs (16) across its entire sequence (Fig. 3) as suggested previously. TAF-I and yeast NAP-I and TAF-I and mouse NAP-I are 24.9% and 18.2% identical with additional conservative substitutions, respectively, while yeast NAP-I and mouse NAP-I are 28.8% identical. Interestingly, TAF-I shows a high degree of homology to the conserved regions between yeast and mouse NAP-I proteins (boxed regions in Fig. 3). The acidic region of TAF-I has some degree of similarity to the acidic portions of a variety of proteins. One example is shown in Fig. 3. Although the detailed mechanism by which TAF-I activates the Ad core DNA replication system is unknown, it is assumed that TAF-I interacts with viral basic core proteins and induces structural change in the Ad core. NAP is known to interact with cellular histones (16). Therefore, we examined the effect of NAP on the Ad core DNA replication. Fig. 4 shows that the relatively high dose of mouse NAP-I stimulated the Ad core DNA replication reaction. TAF-I could also be substituted by yeast NAP-I (data not shown). Currently, it is

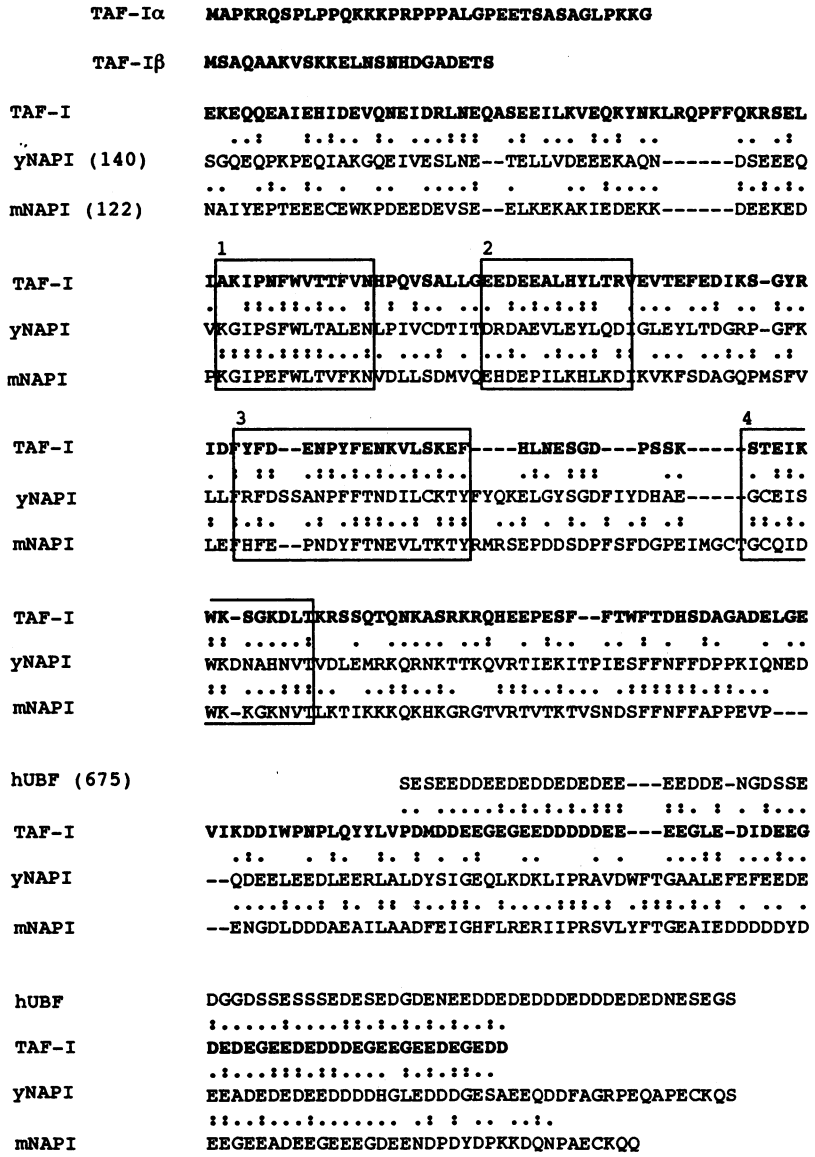

FIG. 3. One possible sequence alignment of human TAF-I, yeast NAP-I, mouse NAP-I, and human UBF. Boldface letters represent the entire amino acid sequences of TAF-I/SET determined by this study and by von Lindern *et al.* (11). The other sequences were obtained from the Protein Identification Resource data bank. Sequences corresponding to the amino acid sequences from position 140 to the C terminus of yeast NAP-I and from position 122 to the C terminus of mouse NAP-I are presented. For human nucleolar transcription factor UBF, which is presented as a representative protein containing an acidic stretch, only the acidic portion of amino acid sequences from position 675 is shown. Pairwise comparison is shown by the symbols : and . for identical and conservative amino acid substitutions, respectively. The four boxed sequences are the regions that are highly conserved between yeast and mouse NAP-I proteins and homologous to TAF-I.

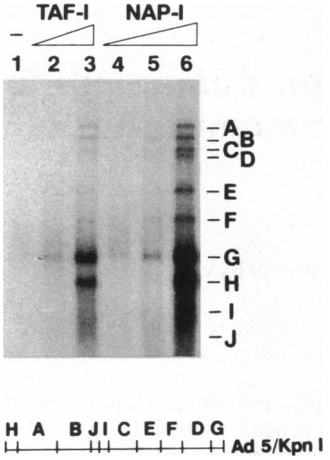

FIG. 4. Effect of mouse NAP-I protein on replication of Ad core DNA. The Ad core DNA replication assay was carried out as described with factors required for replication of naked Ad genome in the absence (lane 1) or presence of 0.01 μg (lane 2) or 0.03 μg (lane 3) of native TAF-I purified from HeLa cells, or 0.1 μg (lane 3), 0.3 μg (lane 4), or 1 μg (lane 5) of mouse NAP-I. *Kpn* I fragments of Ad type 5 DNA are indicated on the right.

unclear whether the acidic region of NAP-I is involved in stimulation of the Ad core DNA replication as is the acidic region of TAF-I. If stimulation of the Ad core DNA replication by TAF-I is due to the dissociation of basic core proteins from DNA, it is possible that NAP-I supports the same reaction in the Ad core DNA replication.

Cellular functions of the protein SET are unknown. *In vivo*, SET is phosphorylated at serine residues and localized predominantly in the nucleus (14). The expression of SET is ubiquitous in a variety of human cell lines and it is thought that SET plays a rather basic function in the organism (11, 14). Recently, the rat homologues of the human *SET* gene, rat *Setα* and -β, both of which may be equivalents of human TAF-Iα and -Iβ, were identified from a rat neonatal kidney cDNA library (17). The rat *Set* gene is highly expressed in developing nephron structures and may play a role in early nephron morphogenesis (17).

The study presented here has revealed that TAF-I required for the Ad core DNA replication is encoded by the *set* gene and its acidic tail is involved in TAF-I activity. It is unclear whether this acidic region is essential for the leukemogenic function of the SET–CAN fusion protein. If this is the case, studies on the mechanism of stimulation of the Ad core DNA replication by TAF-I may contribute to an understanding of the intrinsic function of SET and the leukemogenic function of the SET–CAN fusion protein. Our present hypothesis is that TAF-I interacts with core proteins and facilitates their dissociation from DNA, thereby rendering the DNA accessible for binding of replication proteins. It is interesting to note that the acidic proteins are involved in stimulation of replication and transcription of paramyxovirus genomes (18). We suggest "the acidic molecular chaperones" for these proteins including NAP-I and TAF-I as well as such proteins. It was reported (19) that CAN is a nuclear pore complex protein, nup214 (nucleoporin of 214 kDa), and could play a role in the control of nucleocytoplasmic traffic. It is possible that CAN, fused with SET as the acidic chaperone, is changed in its traffic specificity and/or traffic efficiency.

This work was supported in part by a grant for the Biodesign Research Program from RIKEN (Institute of Physical and Chemical Research, Japan) and a grant-in-aid from the Ministry of Education, Science and Culture of Japan.

1. Hay, R. T. & Russell, W. C. (1989) *Biochem. J.* **258,** 3–16.
2. Stillman, B. W. (1991) *Annu. Rev. Cell Biol.* **5,** 197–245.
3. Salas, M. (1991) *Annu. Rev. Biochem.* **60,** 39–71.
4. Nagata, K., Guggenheimer, R. A., Enomoto, T., Lichy, J. H. & Hurwitz, J. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 6438–6442.
5. Pruijin, G. J. M., van Driel, W. & van der Vliet, P. C. (1986) *Nature (London)* **322,** 656–659.
6. Nagata, K., Guggenheimer, R. A. & Hurwitz, J. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 6177–6181.
7. Brown, D. T., Westphal, M., Burlingham, B. T., Winterhoff, U. & Doerfler, W. (1975) *J. Virol.* **16,** 366–387.
8. Chatterjee, P. K., Vayda, M. E. & Flint, S. J. (1986) *EMBO J.* **5,** 1633–1644.
9. Leith, I. R., Hay, R. T. & Russell, W. C. (1989) *J. Gen. Virol.* **70,** 3235–3248.
10. Matsumoto, K., Nagata, K., Ui, M. & Hanaoka, F. (1993) *J. Biol. Chem.* **268,** 10582–10587.
11. von Lindern, M., van Baal, S., Wiegant, J., Raap, A., Hagemeijer, A. & Grosveld, G. (1992) *Mol. Cell. Biol.* **12,** 3346–3355.
12. Fujii-Nakata, T., Ishimi, Y., Okuda, A. & Kikuchi, A. (1992) *J. Biol. Chem.* **267,** 20980–20986.
13. Hager, D. A. & Burgess, R. R. (1980) *Anal. Biochem.* **109,** 76–86.
14. Adachi, Y., Pavlakis, G. N. & Copeland, T. D. (1994) *J. Biol. Chem.* **269,** 2258–2262.
15. von Lindern, M., Fornerod, M., van Baal, S., Jaegle, M., de Wit, T., Buijs, A. & Grosveld, G. (1992) *Mol. Cell. Biol.* **12,** 1687–1697.
16. Ishimi, Y. & Kikuchi, A. (1991) *J. Biol. Chem.* **266,** 7025–7029.
17. Kim, E.-G., Choi, M. E. & Ballermann, B. (1994) *J. Am. Physiol.* **266,** 155–161.
18. Ishihama, A. & Nagata, K. (1988) *CRC Crit. Rev. Biochem.* **23,** 27–76.
19. Kraemer, D., Wozniak, R. W., Blobel, G. & Radu, A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 1519–1523.
20. Nagata, K., Guggenheimer, R. A. & Hurwitz, J. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 4266–4270.