The SAF-box domain of chromatin protein DEK

Friederike Böhm, Ferdinand Kappes, Ingo Scholten, Nicole Richter, Hiroshi Matsuo, Rolf Knippers and Tanja Waldmann*

Department of Biology, University of Konstanz, 78457 Konstanz, Germany and Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN 55455, USA

Received November 15, 2004; Revised and Accepted January 31, 2005

ABSTRACT
DEK is an abundant chromatin protein in metazoans reaching copy numbers of several millions/nucleus. Previous work has shown that human DEK, a protein of 375 amino acids, has two functional DNA-binding domains, of which one resides in a central part of the molecule and contains sequences corresponding to the scaffold attachment factor-box (SAF-box) domain as found in a growing number of nuclear proteins. Isolated SAF-box peptides (amino acids 137–187) bind weakly to DNA in solution, but when many SAF-box peptides are brought into close proximity on the surface of Sephadex beads, cooperative effects lead to a high affinity to DNA. Furthermore, a peptide (amino acids 87–187) that includes a sequence on the N-terminal side of the SAF-box binds efficiently to DNA. This peptide prefers four-way junction DNA over straight DNA and induces supercoils in relaxed circular DNA just like the full-length DEK. Interestingly, however, the 87–187 amino acid peptide introduces negative supercoils in contrast to the full-length DEK, which is known to introduce positive supercoils. We found that two adjacent regions (amino acids 68–87 and 187–250) are necessary for the formation of positive supercoils. Our data contribute to the ongoing characterization of the abundant and ubiquitous DEK chromatin protein.

INTRODUCTION
The human DEK protein was discovered as a fusion with a nuclear pore protein in a subset of patients with acute myeloid leukemia (1). It was also identified as an autoantigen in a nuclear pore protein in a subset of patients with acute myeloid leukemia (1). It was also identified as an autoantigen in a subset of patients with acute myeloid leukemia (1). It was also identified as an autoantigen in a subset of patients with acute myeloid leukemia (1). In addition, DEK mRNA levels are higher in transcriptionally active and proliferating cells than in resting cells, and elevated mRNA levels are found in several transformed and cancer cells (6,7).

Human DEK is an abundant nuclear protein of 375 amino acids that occurs in copy numbers of more than a million/nucleus. Most nuclear DEK protein is bound to chromatin throughout the cell cycle (8). Its interactions with positive and negative transcription factors (9,10) suggest that DEK may have a role in the transcription complex formation at promoters and enhancers [reviewed in (11)].

Work with the isolated and the recombinant human protein has shown that DEK has an intrinsic DNA-binding activity preferring four-way junction and superhelical over straight DNA and introducing positive supercoils into relaxed circular DNA (12,13). More detailed analysis revealed that DEK has two DNA-binding domains. A first domain is located in a central peptide that includes a conserved sequence element, the SAF (scaffold attachment factor) or SAP (after SAF-A/B; acinus; Pias) box. A second DNA-binding domain is in the C-terminal region of DEK and overlaps with most of the identified phosphorylation sites (Figure 1A). In fact, the DNA-binding properties of DEK are clearly influenced by phosphorylation as phosphorylated DEK binds weaker to DNA than unmodified DEK and induces the formation of DEK multimers (14,15).

The SAF or SAP box is a conserved 35 amino acid motif that contains a number of conserved hydrophobic and charged amino acids, which form two amphipathic $\alpha$-helices that resemble helix 1 and helix 2 of the homeodomain (16,17) (Figure 1B). The SAF box occurs in a number of other chromatin proteins including the matrix or scaffold attachment factors SAF-A and SAF-B, the transcriptional repressor PIAS as well as DNA repair proteins, such as Ku70, the poly(ADP-ribose) polymerase and the fission yeast cruciform-cutting endonuclease (SpCCE) (16). It has been experimentally shown for a few cases that the SAF box can serve as a DNA-binding module.
Here, we investigate DEK-derived peptides that contain the SAF-box motif. We show that the SAF box alone (amino acids 137–187) (Figure 1B) appears not to interact with DNA in solution. However, when many SAF boxes are brought into close proximity, cooperative effects lead to high affinity for DNA. A peptide with amino acids 87–187 (including a sequence of 50 amino acids on the N-terminal side of the SAF box) binds to DNA much like the intact DEK preferring four-way junction DNA over straight DNA. This peptide forms large aggregates in the presence of DNA and is also able to introduce supercoils into relaxed circular DNA. Interestingly, however, the 87–187 amino acid peptide induces negative rather than positive supercoils (as the full-length DEK does). Relatively short stretches of adjacent amino acid sequences are required for the introduction of positive supercoils.

Our experiments contribute to current efforts to better understand the ubiquitous and abundant, yet still enigmatic, chromatin protein DEK.

**MATERIALS AND METHODS**

**Cloning and expression of recombinant DEK fragments**

Cloning and expression of his-tagged DEK full-length as well as peptides 1–250, 1–187, 87–187, 187–375 and Δ87–187 were performed as described previously (15). The fragments 78–208 and 68–226 were cloned and expressed in Escherichia coli. The fragments 87–250, 137–187 and 87–138 with 5′ EcoRI and 3′ XhoI cutting sites were generated by PCR from DEK cDNA cloned in pRSET-A (15). A stop codon was introduced before the XhoI site. The PCR products were subcloned by topo cloning into the pCR Blunt II TOPO vector according to the manufacturer’s protocol (Invitrogen). The cDNA of the DEK fragments were digested with EcoRI and XhoI and then ligated with the linearized pRSET-A vector.

For protein expression, the vectors were transformed in E.coli BL2A (DE3) Lys S. The purification of the his-tagged DEK fragments was performed by Ni-NTA Agarose according to the manufacturer’s protocol (Qiagen) and as described previously (15).

**Topology assay and 2D gel electrophoresis**

Topology assays with the different DEK fragments were performed exactly as described previously (15). A stop codon was introduced before the XhoI site. The PCR products were subcloned by topo cloning into the pCR Blunt II TOPO vector according to the manufacturer’s protocol (Invitrogen). The cDNA of the DEK fragments were digested with EcoRI and XhoI and then ligated with the linearized pRSET-A vector.

For protein expression, the vectors were transformed in E.coli BL2A (DE3) Lys S. The purification of the his-tagged DEK fragments was performed by Ni-NTA Agarose according to the manufacturer’s protocol (Qiagen) and as described previously (15).
Duplex and four-way junction DNA

The preparation of duplex and four-way junction DNA has been described previously (13).

The annealed oligonucleotides were radiolabeled with [γ-32P]ATP using the T4 polynucleotide kinase. Labeled oligonucleotides were incubated with DEK or its derivates and separated on 8% native polyacrylamid gels in 0.5x TBE and visualized by autoradiography.

Electrophoretic mobility shift assay (EMSA) with SV40 DNA

Purified DEK fragments were treated as described for the topology assay. Aliquots containing 0.014 pmol of linearized SV40 DNA were incubated with increasing amounts of DEK as indicated in the Figure 2 legend. The reactions were performed in a total volume of 50 µl and loaded directly onto a 0.6% agarose gel in 0.5x TBE (50 mM Tris-borate, 1 mM EDTA, adjusted to pH 7.8 with boric acid) and run at 2 V/cm overnight. The nucleoprotein complexes were analyzed by SybrGold staining (MobiTec).

Pull down DNA-binding assays

Aliquots containing 10 µl of settled Protein-A Sepharose beads were incubated with 6 µg of monospecific polyclonal DEK antibodies for 1 h at 4°C in nE450 buffer (20 mM HEPES, pH 7.6, 450 mM NaCl, 10 mM sodium bisulfite and 1 mM EDTA) in the presence of 1 µg/µl BSA (New England Biolabs) + 0.5% NP40. The beads were washed twice with 1 ml nE450. The fragments 87–138 and 138–187 were dialyzed against nE100 buffer (20 mM HEPES, pH 7.6, 300 mM NaCl, 10 mM sodium bisulfite, 1 mM EDTA and 1 µg/µl BSA) overnight and incubated at 200 ng polypeptide/200 µl nE450 for 1 h at 4°C with the coupled Protein-A Sepharose beads. The beads were then washed three times with nE100 buffer (20 mM HEPES, pH 7.6, 100 mM NaCl, 10 mM sodium bisulfite and 1 mM EDTA) and 30 ng of radiolabeled and double-digested pMII plasmid (17,18) were added. Competition was carried out with unlabeled,MspI-digested pBlueBacHis2A-vector DNA (fragments ranging between 50 and 1500 bp). After 1 h incubation at 4°C, the beads were washed three times with nE100 and then eluted three times with 100 µl 2% SDS. The eluates were pooled and precipitated by the Wessel–Flügge method (19). The first supernatant was precipitated with 3 M sodium acetate for DNA analysis. The DNA was separated on a 0.6% agarose gel in 0.5x TBE at 2 V/cm and analyzed by autoradiography. Proteins were loaded on an 18% SDS–PAGE and analyzed by immunoblotting.

For the investigation of the full-length DEK, a DNA precipitation assay was used. An aliquot of 600 ng of recombinant DEK (14,15) was incubated in a total volume of 30 µl buffer P (10 mM Tris–HCl, pH 8, 80 mM NaCl and 1 mM MgCl2) in the presence of radiolabeled DNA and competitor DNA for 15 min at room temperature. The resulting protein–DNA aggregates were centrifuged for 15 min at 14000 r.p.m. and the supernatant was collected. The pellets were resolved in 2% SDS and both fractions were analyzed on an agarose gel as described above. For the investigation of a preferential binding of fragments 1–187 and 68–226 to different topological DNA forms, we used the DNA precipitation assay as just described. DNA substrate was partially relaxed SV40 DNA. The assay was performed in a total volume of 30 µl buffer P with 20 ng of DNA and increasing amounts (12.5, 25, 50, 100 and 200 ng) of dialyzed (against nE300) DEK-fragments. After incubation for 15 min at room temperature, the samples were centrifuged for 15 min at 14000 r.p.m. The pellets were dissolved in 1% SDS and deproteinized with Proteinase K for 30 min at 55°C. The DNA was precipitated by ethanol, dried and finally resolved in loading buffer. Final analysis was carried out by agarose gel electrophoresis (0.8%, 17 h, 0.5x TBE, 2 V/cm) and SybrGold staining (MobiTec).

RESULTS

The isolated SAF box binds to DNA

To investigate the DNA-binding properties of the SAF box in DEK, we expressed a peptide comprising amino acids 137–187, the SAF-box peptide (Figure 1B), and compared it with the larger central peptide from amino acids 87 to 187 that we have studied previously (14) and with the full-length recombinant DEK (12,13,15) (Figure 2). We first used EMASs to determine the interaction of these peptides with linearized SV40 DNA. We found that both full-length DEK and the central peptide (amino acids 87–187) induced the formation of large DNA–protein complexes that cannot migrate into the gel (Figure 2B) (14). A minor portion of the protein–DNA complexes carried submaximal amounts of protein and, therefore, migrated into the gel (Figure 2B, DEK lanes 3–6; 87–187 lanes 9–11). In contrast, the SAF-box peptide was unable to bind to DNA under these conditions (Figure 2B, lanes 17–21). An explanation would be that the amino acid sequence which precedes the SAF box and which is present in the central peptide, but absent in the SAF-box peptide, may be responsible for DNA binding. But this could be excluded because a peptide from residues 87 to 138 failed to bind to DNA (Figure 2B, lanes 12–16).

A topology assay is shown in Figure 2C. Both, the full-length DEK and the central peptide could change the topology of closed circular DNA, while the SAF-box peptide was unable to do so (as expected from Figure 2B) (14).

The inability of the SAF-box peptide to bind to DNA was surprising, as the SAF-box motif is assumed to be a major DNA-binding domain. However, we note that Kipp et al. (17) studying an isolated SAF-box peptide of the SAF-A obtained similar results. These authors observed efficient and specific DNA binding when many SAF-box elements were immobilized at closely adjacent sites (17), a phenomenon described as a ‘mass binding mechanism’ that allows many weak DNA-binding proteins to gain efficiency and specificity when they together associate with DNA (20). To determine whether this could also be the case for DEK’s SAF-box peptide, we first loaded Sephadex–protein A beads with DEK-specific antibodies and then added SAF-box peptides that were effectively captured (Figure 3A) and thereby placed in close proximity to each other. The beads were then incubated with labeled DNA fragments derived from a plasmid with a cloned human matrix attachment region (MAR) (17,18). We determined that the immobilized SAF-box peptides efficiently retained the labeled DNA. Competition with unlabeled restricted plasmid DNA revealed a slight, but distinct preference for the MAR fragment MII over the pUC18 plasmid fragment.
Since a preference of DEK for an AT-rich MAR-sequence has not been determined earlier, we repeated the experiment using the full-length DEK instead of the SAF-box peptide. Competition experiments revealed again a preferred binding to the MAR-DNA fragment (Figure 3C). We note though that this preference is less pronounced than that of the SAF box in SAF-A as described by Kipp et al. (17), and it is quite possible that DEK recognizes not so much the sequence, but the unusual DNA structures that are known to occur in DNA with long sequences of AT base pairs (21,22).

In conclusion, we show that the isolated SAF-box peptide of the DEK protein possesses an intrinsic DNA-binding activity that could, however, not be detected in solution, but became apparent after the peptides were immobilized and densely packed on a Sephadex-bead matrix. A central peptide (amino acids 87–187) that includes the SAF box plus N-terminal residues binds as efficiently as the full-length DEK to DNA. Therefore, the sequence preceding the SAF box appears to induce an interaction between single peptides creating a platform that favors a ‘mass binding mode’. Indeed, when incubated with DNA, the central peptide formed large complexes that remained in the slots of the agarose gel exactly as the full-length DEK does (Figure 2B). We investigated whether the central peptide shares other properties with intact DEK.
The central peptide binds to four-way junction DNA

We have recently shown that the full-length DEK preferentially binds to four-way junction DNA (13), and we investigate here whether the central peptide and other selected DEK-derived peptides share this DNA-binding mode. The DNA substrates were a set of labeled oligonucleotides that could be annealed to form either linear or four-way junction DNA (Figure 4A). The full-length DEK as well as the DEK-derived protein fragments bound readily to both DNA forms as detected by the mobility shifts of the resulting protein–DNA complexes (Figure 4B and C, lanes 2 and 7). Their preference for four-way junction DNA was assayed in competition experiments using unlabeled linear DNA as competitor. As expected, the binding to labeled linear DNA could be easily suppressed by competing with unlabeled linear DNA (Figure 4B and C, lanes 8–10), but the binding to four-way junction DNA remained largely resistant to competing straight DNA in cases when the full-length DEK (Figure 4B, lanes 3–5) and peptides including the sequence between residues 87 and 187 were used as binding partners (Figure 4C, 1–187, 87–187, lanes 3–5, respectively). Peptides that lacked the central DNA-binding domain but contained the C-terminal DNA-binding domain (Figure 1) bound efficiently to DNA (Figure 4C, 187–375 and 87–187), but showed no preference for four-way junction DNA, since competing linear DNA suppressed the binding to both labeled linear and labeled four-way junction DNA (Figure 4C, 187–375 and 87–187).

We thus conclude that it is the central DNA-binding domain, including the SAF box, which is responsible for the specificity of DEK for non-linear four-way junction DNA.

The SAF box containing fragment induces negative supercoils

The full-length DEK protein changes the topology of closed circular DNA by introducing positive supercoils (12). As shown previously (14) and above in Figure 2C, the central fragment is also able to induce topological changes in closed circular DNA. However, the directionality of the supercoils introduced by this fragment has not been determined. To investigate this point, we first incubated DEK-derived peptides with closed circular SV40 DNA in the presence of type I DNA topoisomerase. The DNA samples were analyzed after deproteinization by 2D gel electrophoresis using chloroquine in the second dimension. We confirmed that the full-length DEK induced positive superhelical turns, but were surprised to find that the central peptide (amino acids 87–187) induced negative supercoils (Figure 5A). We investigated a number of other DEK-derived peptides and found that sequences on both, the N-terminal side and on the C-terminal side of the central peptide, are necessary for the formation of positive supercoils. The smallest unit that we have found to be capable of inducing positive supercoils includes the amino acids 68–226. Deletions of 18–20 amino acids on either the C-terminal or the N-terminal side of this unit produced peptides that changed the topology of DNA by introducing negative turns.

Thus, a central peptide including the SAF box can change the topology of closed circular DNA much like the full-length DEK does, but the direction of supercoiling is negative rather than positive as in the case of intact DEK. The reason seems to be that the central peptide lacks adjacent sequences which are
Figure 4. Binding of DEK and DEK-derived peptides to four-way junction (4WJ) and straight DNA. (A) Duplex (asterisk) and four-way junction (4WJ) DNA (double asterisk), for molecular sizes see (13). Oligonucleotides 1–4 are partially complementary to each other and assemble to the indicated four-way junction DNA upon annealing. (B) Full-length his-tagged DEK in bandshift assays. Oligonucleotides were annealed, purified on native polyacrylamid gels and then labeled. Labeled 4WJ DNA (lanes 1–5) and labeled duplex DNA (lanes 6–10) was incubated without (lanes 1 and 6) or with recombinant his-DEK (at molar ratio of DEK/DNA 50 for 4WJ and 100 for duplex DNA) without (lanes 2 and 7) or in the presence of increasing amounts of unspecific, unlabeled duplex competitor DNA (at molar ratios: 10, 100 and 500) for 1 h at 37°C. After separation on native 8% polyacrylamid gels, the bands were visualized by autoradiography. (C) Preferential binding of SAF box carrying fragments to 4WJ DNA. Bandshifts were performed as in (B) (molar ratios for all investigated peptides DEK/DNA: 50 for 4WJ DNA, lanes 2–5; 100 for duplex DNA, lanes 7–10 or 0; lane 1, –DEK). Unspecific duplex competitor DNA was present in molar ratios 10, 100 and 500 or absent (lane +DEK). The used DEK derived fragments are indicated above the figures (1–187, 87–187, 187–375 and Δ87–187) and are schematically represented in (D).
somehow involved in changing the sense of supercoiling from negative to positive.

The data of Figure 5 lead to the prediction that the 68–226 fragment, but not the 1–187 fragment, may reduce the negative superhelicity of minichromosomal SV40 DNA with its 25–28 constrained supercoils. This was indeed the case, and the data of Figure 6A show that the 68–226 fragment [like the full-length DEK, Figure 6A; DEK and see (12)] changes the topology of superhelical SV40 DNA, whereas the 1–187 fragment did not. The data could be explained by a different binding preference of the fragments tested. To investigate this point, we used the DNA precipitation assay with partially relaxed SV40 DNA as a binding substrate and increasing amounts of the 1–187 and the 68–226 fragments.

We detected that the 1–187 fragment bound preferentially to DNA topoisomers with reduced negative superhelicity, whereas the 68–226 fragment also bound to highly supercoiled DNA [like the full-length DEK, Figure 6B and C; see (13)]. Thus, the ability of the 68–226 fragment to associate with highly supercoiled DNA may at least partially explain why it reduces the negative superhelicity of SV40 DNA.

DISCUSSION

The SAF or SAP box is an amino acid motif that is detected in an increasing number of nuclear proteins that are involved in various activities, such as nuclear architecture, transcription, RNA processing and DNA repair (16). Most of these activities require an interaction of the protein with DNA. However, experimental evidence that the SAF-box motif directly participates in DNA binding has so far been obtained only for SAF-A (17), PIAS (23), Ku70 (24) and SpCCE1 (25).

In the SAF-A protein, the SAF box appears to be the only DNA-binding motif. It directs the protein to the scaffold/matrix attachment regions (S/MAR) which are AT-rich stretches of DNA that associate with nuclear matrix preparations (26). In addition, SAF-A has a C-terminal RGG motif as an additional module, which is most likely involved in mRNA processing (27–30).

Interestingly, in other proteins, the SAF-box is not the only DNA-binding domain; it increases the specificity and stability of the interactions. Protein Ku70 together with Ku80 forms a protein ring that encircles DNA at their ends as an important step in the non-homologous end-joining pathway. The SAF domain is located in an extreme C-terminal region of Ku70, outside the major DNA contacts, but could function as a barrier to inward movement of Ku70 bound to DNA ends (24).

The SAF box of SpCCE clearly stabilizes the protein’s interaction with four-way junction DNA, but is not its major DNA-binding determinant. However, SAF box is involved in DNA binding as demonstrated by specific amino acid exchanges that reduce the half-life of contacts between SpCCE and DNA. The crucial residues are two basic amino acids close to the second helix in the helix–loop–helix domain that constitute the SAF box (25) (Figure 1).
Like SpCCE and Ku70, DEK has two DNA-binding domains, namely a central domain including the SAF-box and a C-terminal domain of as yet undetermined structure (Figure 1A). However, it appears that the SAF-box domain is the main DNA-binding activity because a central fragment that includes the SAF box, but not the C-terminal binding domain, has DNA-binding properties that are similar to those of the entire DEK molecule. More precisely, DEK and the

Figure 6. The binding of DEK-derived polypeptides to superhelical DNA. (A) Topology assay using SV40 minichromosomes. Left panel, control DNA. Full-length DEK, the 1–187 fragment and the 68–226 fragment were incubated at molar ratios of 50, 100 and 150 protein/DNA for 1 h at 37°C in the presence of topoisomerase I as described previously (12). The DNA was purified and analyzed by agarose gel electrophoresis and SybrGold staining. (B) DNA precipitation assay. Partially relaxed SV40 DNA (lane input, 20 ng) was incubated with increasing amounts of the polypeptides as indicated. After incubation for 15 min at room temperature, the protein–DNA complexes were precipitated. The DNA was purified and analyzed by agarose gel electrophoresis and SybrGold staining. (C) Densitometric analysis. The intensities of a selected DNA topoisomer with low superhelical density (filled square) and the intensities of highly supercoiled form I DNA (filled triangle) were determined by the densitometric program AIDA (Raytest). The results were plotted as relative amounts of precipitated DNA versus amounts of protein compared with input DNA.
central domain prefer four-way junction over straight B-form DNA and induce supercoils into relaxed circular DNA. Peptides including the C-terminal domain bind well to DNA, but do not have the specificities of the full-length DEK or the central peptide. We assume that the second DNA-binding domain stabilizes the interaction of DEK with DNA and modulates its interaction in vivo and in vitro as it overlaps with major phosphorylation sites. Indeed, previous work has shown that a phosphorylation of DEK weakens DEK–DNA contacts (15).

Like SAF-A, DEK’s isolated SAF box has a very low affinity to DNA that cannot be determined by standard DNA-binding assays in solution (Figure 2B). Rather, it is necessary to fix many SAF boxes on a Sephadex matrix forming a DNA-binding platform that firmly associates with DNA. The combined binding activities of immobilized SAF boxes show a slight preference for AT-rich S/MAR sequences (Figure 3B), but it is in any case much less pronounced than the affinity of SAF-A for S/MAR (17). It is therefore not clear whether DEK’s preference for S/MAR is physiologically significant. In addition, AT-rich sequences are known to assume unusual DNA structures (21,22) and it is possible that DEK recognizes these structures and not the sequence. We note though that one-tenth or so of all nuclear DEK fractionates with standard nuclear matrix preparations (31,32) (data not shown). However, the majority of nuclear DEK can be released from chromatin by mild digestion with micrococcal nuclease (8). It will be an interesting problem for future research to determine how DEK is distributed on chromatin, and how it is directed to its binding sites on chromatin. A possibility is that DEK associates with transcription factors at promoters or enhancers changing the path of the DNA to which it is bound, much like the better known HMG-A or HMG-B architectural proteins (33–36). Like DEK, these proteins are known to preferentially bind to four-way junction and superhelical DNA (36–38).

Another similarity is that HMG-A and HMG-B proteins, such as DEK, induce supercoils into relaxed circular DNA. Whether this property has physiological consequences in vivo is not known. The topological changes are probably caused by the extensive twisting of DNA that occurs in the large DEK–DNA complexes which form in vitro (13). In any case, it is interesting to note that the central peptide (amino acids 87–187) introduces negative supercoils, whereas a slightly larger peptide (amino acids 68–226) induces positive supercoils (like the full-length DEK). This change in directionality cannot simply be explained by the known properties of DEK. Clearly, a determination of the 3D structure would help to successfully address this issue.

In conclusion, we have shown here that the SAF box is the major DNA-binding domain of DEK because a central fragment of ~100 amino acids that includes the SAF box has many of the DNA properties of the full-length DEK, while the C-terminal second DNA-binding domain appears to influence the overall protein–DNA contacts.

ACKNOWLEDGEMENTS

This work was supported by Deutsche Forschungsgemeinschaft. Funding to pay the Open Access publication charges for this article was provided by Deutsche Forschungsgemeinschaft.

REFERENCES