Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 6892–6897, July 1996
Biochemistry

Mutational analysis of NM23-H2/NDP kinase identifies the structural domains critical to recognition of a c-*myc* regulatory element

(PuF/transcription factor/metastasis)

EDITH H. POSTEL*, VALERIE H. WEISS†, JUTTA BENEKEN‡, AND AJAY KIRTANE§

Department of Molecular Biology, Princeton University, Princeton, NJ 08544-1014

Communicated by Arnold J. Levine, Princeton University, Princeton, NJ, February 20, 1996 (received for review December 6, 1995)

ABSTRACT **NM23-H2, a presumed regulator of tumor metastasis in humans, is a hexameric protein with both enzymatic (NDP kinase) and regulatory (transcriptional activation) activity. While the structure and catalytic mechanisms have been well characterized, the mode of DNA binding is not known. We examined this latter function in a site-directed mutational study and identified residues and domains essential for the recognition of a c-*myc* regulatory sequence. Three amino acids, Arg-34, Asn-69, and Lys-135, were found among 30 possibilities to be critical for DNA binding. Two of these, Asn-69 and Lys-135, are not conserved between NM23 variants differing in DNA-binding potential, suggesting that DNA recognition resides partly in nonconserved amino acids. All three DNA-binding defective mutant proteins are active enzymatically and appear to be stable hexamers, suggesting that they perform at the level of DNA recognition and that separate functional domains exist for enzyme catalysis and DNA binding. In the context of the known crystal structure of NM23-H2, the DNA-binding residues are located within distinct structural motifs in the monomer, which are exposed to the surface near the 2-fold axis of adjacent subunits in the hexamer. These findings are explained by a model in which NM23-H2 binds DNA with a combinatorial surface consisting of the "outer" face of the dimer. Chemical crosslinking data support a dimeric DNA-binding mode by NM23-H2.**

Altered *nm23* gene expression is associated with a multitude of phenotypes including inhibition of tumor metastasis (1–5), oncogenesis (6, 7), cellular proliferation (8, 9), development, differentiation (10–14), and apoptosis (14). Although *nm23* was discovered a decade ago, there are still no molecular or biochemical explanations for these data. The nature of these observations suggest, however, that NM23 is involved in the regulation of other genes important to cell growth and differentiation. Indeed, such a role for human NM23-H2 has been recognized in c-*myc* gene transcription (15–17).

nm23 genes encode the 17-kDa subunits of nucleoside diphosphate kinases (NDPKs), housekeeping enzymes that catalyze the transfer of γ-phosphates between nucleoside tri- and diphosphates (18). The *nm23* gene family is large and highly conserved among species (19). Human *nm23-H1* (1) and *nm23-H2* (3) are 88% homologous and are closely linked on chromosome 17q21 near the BRCA-1 gene locus (20). The NM23-H1 protein is more closely associated with metastasis inhibition and signal transduction (1, 2, 21), has an acidic pI, and is also known as NDPK-A (22). The second variant, NM23-H2 (3), is a basic protein identical to NDPK-B (22) and to the human PuF factor, a transcriptional activator of the c-*myc* protooncogene (23, 24). A third gene located on chromosome 16, known as *DR-nm23,* is 70% identical to *H1* and *H2* and may play a role in normal hematopoiesis and in the induction of apoptosis (14). Overexpression of all three human *nm23* genes is postulated to contribute to differentiation arrest.

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.

The 3-dimensional structures of several NM23/NDPKs are known and the catalytic mechanisms have been well characterized (25–29). All eukaryotic NM23/NDPKs are hexameric, a conformation which is required for stability as well as for efficient enzyme catalysis (25–29). In contrast, the biochemical and molecular bases of the of NM23 regulatory functions are not known. As the developmental and the metastatic functions of NM23 appear to be independent of the NDPK enzymatic activity (2, 11, 30, 31), it seems likely that at least some of the biological properties are consequences of the transcriptional function, particularly because one target of NM23-H2 is the c-*myc* gene, itself a regulator of cell proliferation and differentiation (8, 15, 17, 32–34). This notion is partly supported by the finding that DNA binding and transcriptional activation, at least *in vitro*, can occur in the absence of the phosphotransferase activity (35).

To attempt a fuller understanding of the NM23 function, we have undertaken a study of its transcriptional activity. As a starting point for this analysis, we examined in a previous study (35) the relevance of the NDPK activity to the DNA-binding function and concluded that, whereas a missense substitution of the catalytic residue (His-118 with phenylalanine) inactivated the phosphotransferase function, the mutation had no effect on DNA binding or *in vitro* transcription. Thus, with a single point mutation we essentially have separated the catalytic and the DNA-binding functions, although the physiological consequences of this finding are not known yet. Here we used a systematic site-directed mutational analysis to identify residues and domains of NM23-H2 that are involved in DNA binding. On the basis of these analyses, a model for the NM23-H2 DNA-binding domain is presented.

MATERIALS AND METHODS

Mutagenesis. Point substitution mutations were introduced into *nm23-H2* cDNA by the Unique Site Elimination method as described (35). The mutagenic primers were 27–39 bases long and had a greater than 50% GC content. The sequence of the oligonucleotides used to generate the DNA-binding defective mutants were, with base changes indicated in boldface letters, as follows: 5′-gagcagaagggattc**gcc**ctcgtggccatgaag-3′ (R34A); 5′-gcagaagggattc**gac**ctcgtggccatg-3′; (R34D); 5-′cagaagggattc**gg**cctcgtggccatg-3′ (R34G); 5′- gggctggtgaagtacatg**cac**tcagggccggttgtggcc-3′ (N69H); 5′-cagcctatggttt**cac**cctgaagaactggttgac-3′ (K135H).

Protein Production and DNA Binding. Mutant and wild-type proteins were overproduced in *Escherichia coli* and purified by ammonium sulfate fractionation and hydroxylapatite chromatography as described (15). DNA binding was assessed in gel

Abbreviation: NDPK, nucleoside diphosphate kinase.

*To whom reprint requests should be addressed.
†Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115.
‡Present address: Program in Biochemistry, Cell and Molecular Biology, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205-2185.
§Present address: College of Physicians and Surgeons, Columbia University, New York, NY 10032.

FIG. 1. Location of amino acid substitution mutations in the primary sequence of NM23-H2/NDPK. (*A*) Schematic representation of NM23-H2 (hatched box) indicating major secondary structural features. The "head," the Kpn loop, and the C- terminal segment are exposed to the outer surface in the hexamer (25–29). (*B*) Partial amino acid sequence of NM23-H2 showing the mutated residues in boldface type.

electrophoretic mobility shift assays (15, 24, 35). Standard reactions were carried out with [^{32}P]-end-labeled c-*myc* fragments containing the −164 to −110 nuclease-hypersensitive element (NHE), the known site of interaction with PuF/NM23-H2 (15, 24). One fragment was 105 bp long and generated by PCR (15, 35); the other was a 34 bp double-stranded synthetic oligonucleotide with the sequence 5′-ctccccaccttccccaccctccccaccctcccca-3′. The two DNA fragments provided qualitatively similar results although, as noted before (24), higher affinities were observed with the longer probe. Stoichiometric amounts of DNA and protein (generally 1 ng probe: 200-2000 ng of hexameric protein) were mixed together in 10-μl reactions also containing 100 ng poly(dA–dT) in 0.1 M HM buffer (20 mM Hepes, pH 7.9/5 mM $MgCl_2$/0.1 mM EDTA/0.1 M KCl/1 mM DTT/20% glycerol/protease inhibitors; refs. 15 and 24) and incubated 20 min on ice. The NM23/DNA complexes were resolved on 5% native polyacrylamide gels and detected by autoradiography as described (15, 35).

Size Exclusion Chromatography. The molecular size of wild-type NM23-H2 was initially determined using a Pharmacia FPLC Superose 12 column (29 × 1.5 cm), equilibrated with 0.1 HM buffer and calibrated several times with low and high molecular weight protein standards (Pharmacia). A standard curve was constructed by plotting K_{av} versus log M_r. K_{av} was calculated from the elution volumes V_e, from total column volume V_t, and from the void volume V_0 using the equation $K_{av} = V_e - V_0/V_t - V_0$. The void volume ($V_0$) was determined by measuring the eluted volume of blue dextran. From the FPLC column the molecular weight of recombinant NM23-H2 was determined to be 100,000 ± 10,000. Based on the amino acid sequence (15), the calculated molecular weight of the hexamer is 105,000 (22). In subsequent experiments, a conventional 27 × 1.5 cm Sephacryl S-200 HR (Pharmacia) column was used, from which the same M_r was obtained for the wild-type hexamer. This column was recalibrated frequently and used for the sizing of the mutant proteins.

NDPK Enzyme Activity. NDPK activity was measured spectrophotometrically in a coupled pyruvate kinase–lactate dehydrogenase assay and the specific activities calculated as described (35).

Glutaraldehyde Crosslinking. Wild-type NM23-H2 (200 ng) was incubated with or without DNA in a standard DNA-binding reaction, then treated with glutaraldehyde (Sigma) freshly diluted in water. Crosslinking was allowed to proceed for 30 min at 25°C and was terminated by the addition of 1 μl of 1 M lysine. As indicated, reactions were treated with 1 unit of DNase (Promega) for 10 min at 37°C, after which samples were boiled in SDS sample buffer and electrophoresed in 15% polyacrylamide gels. To detect DNA associated with proteins, the wet gels were exposed to x-ray film at 4°C overnight prior to detection of proteins by immunoblotting. We used mouse polyclonal anti-NM23-H2 antiserum and horseradish peroxidase-conjugated secondary antibodies (Vector Laboratories) as described (15).

RESULTS

Rationales for Designing the Mutations. Because the crystal structure of NM23 has not suggested a mode of DNA binding (25–29), the design of mutations was, within the context of the 3-dimensional structure, based on the biochemical and biological properties of NM23 including: (*i*) the presence of transcription factor motifs in the protein sequence and likely regulatory domains, (*ii*) identification of functional domains using monoclonal antibodies, (*iii*) a 12% sequence diversity between NM23-H1 and H2 and the demonstration that H1 does not bind to c-*myc* promoter DNA, and (*iv*) mutations occurring naturally in *nm23* that are associated with cancer or developmental defects.

Transcription factor motifs as targets of mutagenesis. nm23-H2 encodes a potential bZIP-like motif involving helices α1 and αA (ref. 3; Fig. 1). Consistent with a heterodimerization-based transcriptional activation mechanism is the behavior of endogenous PuF/NM23-H2 in HeLa cell fractions (15) and a Leu → Val substitution at position 48 (in the middle leucine) in an aggressive case of childhood neuroblastoma (7). Although the "basic" region of the putative zipper motif in NM23 is part of helix α1 (Fig. 1), which constitutes the protein interface–oligomerization domain between two adjacent monomers (25–29), we nonetheless mutated these residues (R18, K26, R27, K31, R34, and K39) because of the possibility that DNA may dissociate the interacting helices prior to binding. A similar model has been proposed by Janin and coworkers (26).

Because the αA helix contains the leucines (at positions 41, 48, and 55) and because it is ideally located on the surface of the hexamer (25–29), this helix could be a protein interacting domain. Alternatively, the αA helix could be considered the DNA-recognition helix, for it is rich in charged residues that could stably interact with DNA (e.g., R42, E46, H47, and Q50). In fact, the αA helix in NM23-H2, unlike in H1, has an overall basic charge. Finally, the αA helix could function in a dual capacity by binding both protein and DNA.

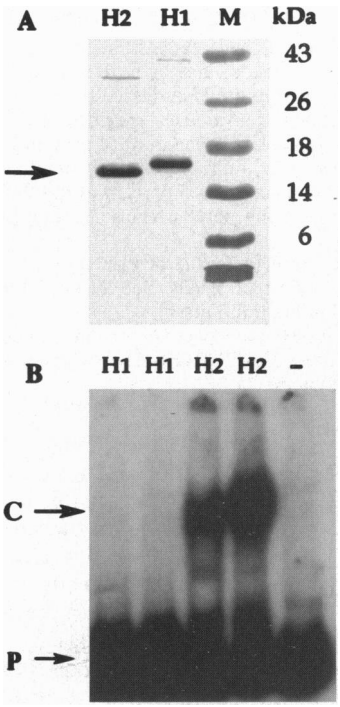

FIG. 2. DNA-binding activity of NM23-H2 and H1. (*A*) SDS/PAGE of purified H1 and H2 proteins used in DNA binding. Migration of molecular size markers is shown on the right in kDa, with arrow on the left pointing to the 17-kDa protein monomers. Although the calculated molecular weights of both proteins are nearly identical, H1 typically migrates with a higher apparent molecular weight than expected. (*B*) Electrophoretic mobility-shift assay showing DNA binding by NM23-H2, but not by H1; 200 and 1000 ng of each protein was tested with the 105-bp fragment. The same result was obtained with the 34-bp probe (not shown). Arrows on left indicate migration of the free probe (P) and the specific complexes formed (C).

Table 1. Properties of NM23-H2/NDPK missense mutants

Mutation	Expression and stability in *E. coli**	DNA binding†	NDPK activity‡	Oligomeriz state§
Wild type	+	+	+	Hexamer
R18G	±	±		
K26G	+	+		
K26Q	−			
R27G	−			
R27Q	+	+		
R27V	+	+		
K26G + R27G	−			
K31G	±	+		
K31A + P101G	±	+	Mixed	Mixed
R34G	+	−	+	
R34A	+	−	+	Hexamer
R34D	+	−	+	
K39A	+	+	+	Hexamer
L41M	+	+	+	
L41M + R42Q	+	+	+	
R42Q	+	+		
L41M + L48V	+	+	+	
L48V	+	+	+	
L48V + L55V	+	+	+	
L41M + L48V + L55V	+	+	+	
S44A	+	+		
E46L	+	+		
H47D	+	+		
Q50E	+	+		
H51F	+	+		
Y52V	±	+	±	Hexamer
Stop69	−			
N69H	+	−	+	Hexamer
E79A	+	+		
E79A + P62A	−			
R88A	+	+		
P96S	+	+		
P101S	+	+	+	
P101G	+	+	+	Hexamer
R105A	+	+		
R114A	+	+		
H118F	+	+	−	Ref. 35
K124E	+	+		
K124E + K135H	−			
K135H	+	−	+	Hexamer
Y142A	+	+		
K143T	+	+	+	Hexamer
H147F	+	+		
D148N	+	+		

*Normal expression is indicated by +, its absence by −. Reduced or abnormal protein expression/and or stability is indicated by ±.

†Approximately normal DNA binding measured as specific activity (intensity of sequence specific shifted complex/amount of protein in electrophoretic mobility-shift assay) is indicated by +; complete absence of DNA binding or a residual activity <5 percent of the WT is indicated by −. All proteins, with the few noted exceptions, showed wild-type levels of DNA binding.

‡NDPK activity was measured in a coupled assay as described (35). +, specific activities comparable to wild-type levels. See Table 2 for specific activity values of DNA-binding mutants.

§As determined by SEC. All proteins examined were determined to be hexameric, with the exception of the K31A/P101G mutant, which was a mixture of different order oligomers.

Another DNA-binding mode is suggested by the presence of a potential HTH motif, comprising the αA–turn–α2 structural domains (Fig. 1). These two helices, referred to as the "head" of the NM23-H2 molecule (29), in fact protrude on the side and, although not properly oriented for DNA binding (25), they could, after minor conformational changes, directly access the DNA substrate. In the HTH model, the αA helix would be considered the probe helix for DNA binding.

Combined functional and immunological approach points to α2/β3 as involved in DNA binding. As a preliminary step to mutagenesis we used monoclonal antibodies raised against NM23-H2 to identify functionally important domains. One monoclonal antibody, mAb3E4, which inhibited DNA binding by both HeLa and recombinant NM23-H2, recognized an epitope consisting of residues 63–79 located to the α2/β3 (E.H.P. and C. Ferrone, unpublished results). This finding further raised interest in the HTH motif and the possibility that helix α2 is involved in DNA recognition, either alone, or as part of another as yet unidentified motif.

Sequence diversity between NM23-H1 and H2 may account for their disparity in DNA binding. The possibility that the closely related NM23 variants H1 and H2 might differ in their DNA-binding potential was suggested by their differences in charge (22), regulation (3), and amino acid sequence of potential regulatory regions, including helix αA and the extended C-terminal fragment (25–29). These assumptions were borne out by experimental evidence. First, we tested purified human erythrocyte NDPK A/H1 and B/H2 enzymes (22) and observed significant differences in terms of their DNA-binding potential to c-*myc* promoter DNA (E.H.P., unpublished results). Second, Hildebrandt et *al.* (36) reported the absence of binding by NM23-H1/NDPK-A to a c-*myc* NHE oligonucleotide (unpublished results). Third, we demonstrate in this report that NM23-H1 does not bind to the −164 to −110 c-*myc* promoter sequence (Fig. 2). We therefore substituted all of the nonconserved H2 residues with the potential for H-bonding with DNA, with H1 amino acids, including R42Q, E46L, H47D, Q50E (αA helix) and K124E, K135H, K143T, H147F, and D148N (C terminus). Interestingly, one of the unconserved residues, Asn-69, located at the 3′ end of the α2 helix, is within the critical DNA-binding epitope recognized by mAb3E4. Asn-69 is replaced by His-69 in NM23-H1, is on the protein surface (26, 28, 29), and is capable of sequence specific contact with DNA (37). Based on these considerations we made the Asn-69 → His-69 substitution a major target of our studies.

Naturally occurring mutations associated with cancer or development may have a DNA-binding defect. Because most of the naturally occurring mutations in the tumor suppressor protein p53 disrupt DNA-binding interactions (38), we might expect that similar mutations in NM23, a putative metastasis suppressor, may also be involved in DNA binding. To date, only a handful of NM23 mutations have been identified that are associated with cancer or abnormal development. One of these, discussed above, is the L48V mutation found in neuroblastoma (7). Another possibility is the conditional lethal *killer of prune* (*kpn*) mutation in Awd, the *Drosophila* NM23 homologue (11). Kpn is a Pro → Ser substitution at position 97 (P96 in NM23-H2), located within a structural motif called the "Kpn loop" (named after the mutation). The Kpn loops are considered critical for trimeric interactions with neighboring subunits and are exposed on the top and bottom surface of the hexamer (25–29). The involvement of the Kpn loops in DNA binding would thus require substantial conformational changes. In addition to P96S, we also targeted residues R88, P101, and R114 in the loop, because of their potential, based on charge, of influencing DNA binding. Another residue of interest was Ser-44, whose phosphorylation has been associated with the metastatic phenotype (39), and, although conserved between H1 and H2, this serine is near the surface of the molecule and could thus readily make contact with DNA.

Properties of Site-Directed Mutations. A total of 30 individual amino acids were targeted for mutagenesis and, including double and triple substitutions, over 40 mutant proteins were examined for DNA-binding activity. Those found relevant or interesting were also examined for NDPK activity and oligomerization state. The mutations (Table 1) can be classified according to the following phenotypes: (*i*) mutations that have no effect on protein stability and activity (the majority), (*ii*) those that affect intrinsic stability and solubility of the protein (six), and (*iii*) mutations that disrupt DNA binding (three). Mutations belonging to category 2 involving residues K31, Y52, and P96 affected solubility (salting out by ammonium sulfate) and/or stability (hexameric conformation) to

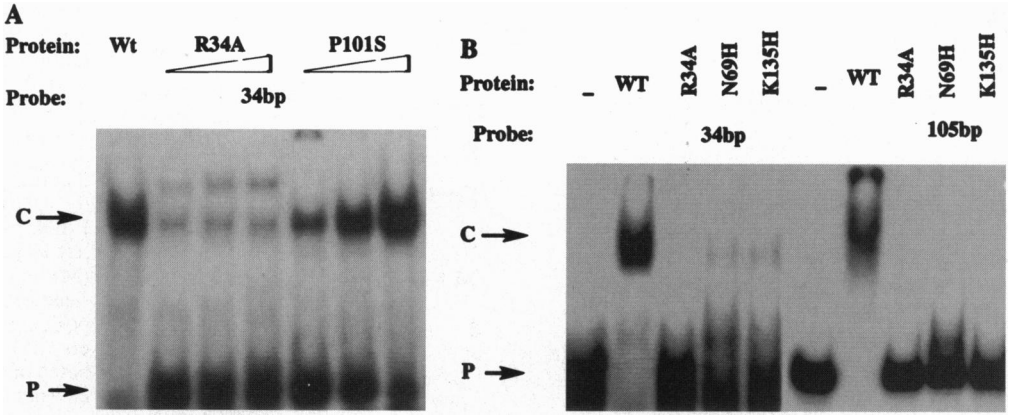

FIG. 3. Electrophoretic mobility-shift assays showing DNA-binding activity of mutant proteins. (*A*) Titration of R34A and P101S mutants showing the R34A DNA-binding defect. Increasing concentrations of mutant R34A and P101S proteins (100, 300, and 1000 ng, respectively), and 500 ng of WT protein were tested for DNA binding to 1 ng of the 34-bp probe. Arrows on left indicate migration of the free probe (P) and the specific complexes formed (C). The lower band (complex C) in the R34A titrations represent the residual binding activity of the mutant (< 5%), whereas the upper band appears to be nonspecific binding that sometimes appears with the 34-bp probe. We have seen this signal with other mutants, as well as with a WT protein preparation that bound DNA poorly (not shown). (*B*) DNA-binding defect of R34A, N69H, and K135H mutant proteins showing less than 5% residual activity. Both the WT and the R34A proteins represent different preparations from those used in *A*. Five hundred nanograms of protein were used in each reaction. Arrows on left indicate migration of the free probe (P) and the specific complexes formed (C).

some degree and were therefore difficult to assess for DNA binding and, although not found defective, they may have reduced activity. One mutant deficient in catalytic activity (H118F), but normal with respect to DNA-binding functions, has already been described (35).

α1-β2-αA mutants. Some of the mutants could not be produced in *E. coli* because the substitutions probably destabilized the protein fold (e.g., K26Q and R27G). When proteins with other substitutions at these positions were produced—e.g., K26G, R27Q, and R27V—they did bind to DNA normally (Table 1), suggesting that residues R18, K26, R27, and K31, and therefore helix α1, could not be involved in DNA binding. These results argue against the validity of the bZIP model and others using α1 as the probe helix for DNA recognition (3, 26). Mutations in Arg-34, which is located on the edge of the β2 sheet, on the other hand, severely impaired DNA binding (Table 1 and Fig. 3). Fig. 3 shows, in a typical electrophoretic mobility-shift assay, the DNA-binding defect of the R34A protein in comparison to another mutant, P101S, and to the wild-type proteins. Clearly, Arg-34 is essential for DNA binding by NM23-H2.

αA turn α2 mutants. We obtained only one mutant from this group, N69H, which is defective in DNA binding (Table 1 and Fig. 3*B*). Asn-69 is at the C terminus of the α2 helix and was expected to be involved in DNA binding, both on the basis of the antibody inhibition experiments and because of the lack of conservation between H1 and H2 (N is substituted by H in NM23-H1). Contrary to expectations, however, the other nonconserved residues, R42, E46, H47, and Q50, that are part of helix αA and the external head of the molecule, do not appear to be involved in DNA binding to this c-*myc* sequence as judged by our electrophoretic technique. While the involvement or subtle influence of these and the other residues apparently uninvolved cannot yet be ruled out, we suggest that αA may be a protein interacting domain instead and that some of these charged residues (through salt bridging), as well as the leucines (through hydrophobic interactions), may be critical to the binding of effector molecules modulating NM23 activity. The overall basic charge of the αA helix suggests the interacting protein(s) is likely to be acidic.

Table 2. Summary of properties of important NM23-H2/NDPK mutants

Mutation	Expression in *E. coli*	DNA binding	NDPK activity*	Oligomerization state†
WT	+	+	+	Hexamer
H118F	+/↓	+	−	Not tested
R34A	+	−	+	Hexamer
N69H	+	−	+	Hexamer
K135H	+	−	+	Hexamer

*Specific activities in units/ml (35) were 105 (WT), 90 (R34A), 84 (N69H), and 83 (K135H).

†Molecular sizes estimated by SEC were: 102,000 (WT), 105,000 (R34A), 97,000 (N69H), and 98,000 (K135H). See Table 1 legend for further details. The H118F mutant (35) is listed for comparison.

Kpn loop and C-terminal mutations. We found that neither the *killer-of-prune* mutation P96S nor any of the other residues in the Kpn loop we changed (P101, R88, and R114) appear to be involved in DNA binding (Table 1 and Fig. 3). Of the C-terminal residues we targeted, only one mutant, K135H, was defective in DNA binding (Fig. 3 and Table 1). This mutation confirms both the expectation that unconserved residues are involved in DNA binding (Lys-135 is substituted with His-135 in NM23-H1) and the importance of the C-terminal segment that wraps around the outer surface of the protein (28, 29). Table 2 summarizes the properties of the DNA-binding defective mutants.

Glutaraldehyde Crosslinking Suggests NM23-H2 Binds DNA as a Dimer. The quaternary structure of NM23/NDPK (22, 28, 29), including recombinant NM23-H2 used in these studies, has been consistently measured as hexameric, a conformation that appears to be required for stability and NDPK activity (26–28). Although the importance of the hexameric conformation to DNA binding is not known yet, in every case we examined, the protein prior to DNA binding was hexameric (Tables 1 and 2). To study the oligomeric structure of NM23-H2 that binds to DNA we performed chemical crosslinking experiments in conjunction with SDS/PAGE gels and Western blotting. Fig. 4 shows that in the abscence of DNA, the uncrosslinked protein under dissociating and denaturing conditions is, as expected, monomeric (apparent size 17 kDa). In the presence of 0.01% glutaraldehyde, however, dimers (apparent size 34 kDa), as well as some higher order oligomers (trimers, tetramers, and hexamers) are also formed. In the presence of DNA, the same protein crosslinking pattern is observed; however, the SDS/PAGE gels exposed to film prior to Western blotting indicate that the [^{32}P]-end-labeled DNA fragment is associated almost entirely with the dimeric subunit (Fig. 4*A*).

DISCUSSION

We identified three amino acids that are absolutely essential for DNA binding by NM23-H2: Arg-34, Asn-69, and Lys-135. Based upon their properties (summarized in Table 2) we conclude: (*i*) amino acids involved in DNA recognition are restricted to regions of the protein that are exposed to the surface of the hexamer; (*ii*) two out of three are residues not conserved between the closely related isotypes NM23-H2 and H1, which are 88% homologous but differ in their DNA-binding potential. This suggests that the DNA-binding specificity of NM23 resides in part with the nonconserved residues, but does not rule out the possibility that NM23-H1 may bind

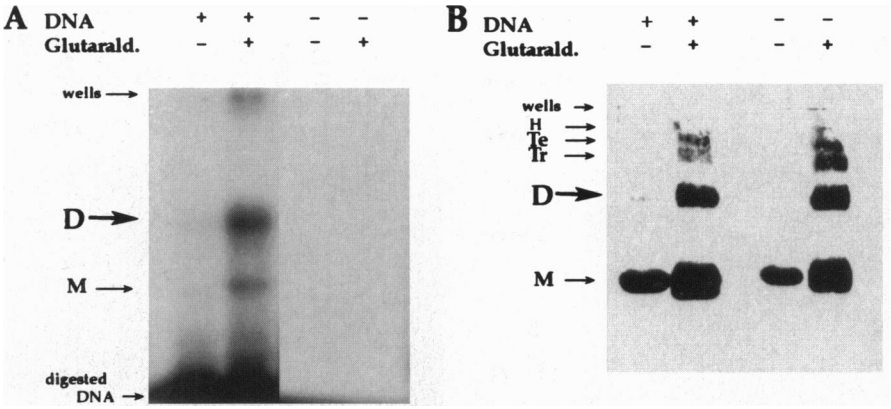

FIG. 4. Glutaraldehyde crosslinking of NM23-H2 to DNA shows dimeric binding. A glutaraldehyde titration experiment (not shown), using concentrations between 0.001–0.1%, indicated that the 0.01% range was optimal; at the higher concentrations of glutaraldehyde NM23-H2 formed artifactual crosslinks that resisted electrophoresis and also degraded the DNA. (*A*) Autoradiograph of wet SDS/PAGE gel prior to immunoblotting. Arrows on left indicate expected positions of monomers (M) and dimers (D). (*B*) Immunoblot of the same gel. Arrows on left indicate the expected positions of monomers (M), dimers (D), trimers (Tr), tetramers (Te), and hexamers (H) based on the migration of molecular size markers (not shown).

to a different DNA sequence specified by the histidines. (*iii*) All three DNA-binding defective mutant proteins are active enzymatically and appear to be stable hexamers, implying that Arg-34, Asn-69, and Lys-135 perform at the level of DNA recognition and not through stabilization of the NM23 fold. Although the Arg-34 side chain is not fully exposed (ref. 28; Fig. 5), it could still contact DNA directly after minor conformational changes such as an induced curvature on the protein, or it could make contact with a protruding structural component of the DNA. Another possibility is that R34 is important to DNA binding for structural reasons, e.g., by holding the C terminus in place, as was suggested by Y. W. Xu and J. Janin (personal communication) on the basis of modelling of Arg-34 side chain interactions with Y142 (C terminus) and E79 (β2). (*iv*) All three mutants function normally as NDP kinases, thus confirming our earlier conclusions, drawn from the properties of the active site H118 mutation, that there exist two separate functional domains for NM23/NDPK: a catalytic domain and a DNA-recognition domain. These data also imply that NDPK activity and DNA binding are not mutually exclusive functions, thus allowing for a possible interaction of the two domains *in vivo* in a capacity that requires both DNA binding and nucleotide phosphorylation, such as might occur during DNA and RNA synthesis. Experiments using cell transfection assays (16) are currently in progress to assess the requirement of the NDPK function and DNA binding in the transactivation of a c-*myc* promoter in the cell.

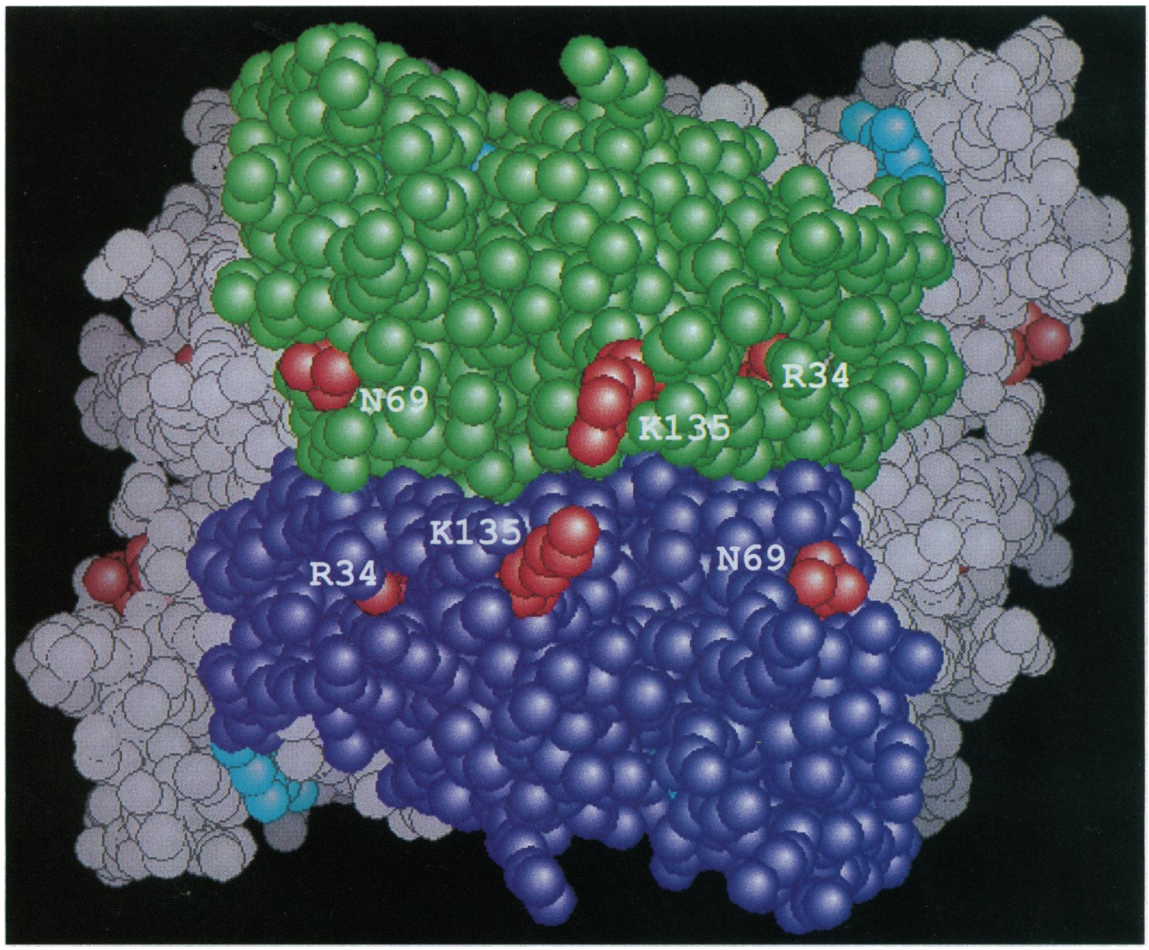

FIG. 5. Space-filling model of the NM23-H2 hexamer. Residues implicated in DNA binding are in red, GDP bound to the NDPK active site in cyan (28). The two subunits of the front dimer are shown in green and purple, the other two dimers are in grey. The model was drawn with the GRASP computer graphics program.

Although each of the residues critical for DNA binding are located on separate structural motifs (Arg-34 at the edge of β2, Asn-69 on the C terminus of the α2 helix, and K135 on the unstructured C-terminal segment; Fig. 1), they line up along the equator on the "outer" face of the hexamer (Fig. 5). A model for these findings is provided on the basis of the known crystal structure (28, 29). The NM23/NDPK hexamer consists of three vertical dimers around a 3-fold axis, or two horizontal trimers along a 2-fold axis. The dimers are held together by two α1 helices that are in antiparallel contact with each other. DNA-binding residues Arg-34, Asn-69, and Lys-135 are clustered around the interface of the dimer (Fig. 5). As our genetic data exclude the dimer interface from participating in DNA binding, we suggest a model in which DNA makes contact with residues of two adjacent monomers near the 2-fold axis and on the outer surface of the hexamer. This creates a combinatorial DNA-binding surface with possibly a single active site composed of residues donated by both monomers of a dimer.

We have previously reported that HeLa PuF/Nm23-H2 makes contact with the major groove of DNA (24). The methylation interference contacts GGGTGGG (repeated 3 times within the c-*myc* promoter fragments used in these studies) are consistent with asymmetric binding and the possibility that some of the DNA-binding residues make contact with opposite sides of the major groove of DNA. The dimensions of the dimer interface (≈30 Å) are compatible with contact by a single turn of B-DNA. However, as the structure of the c-*myc* promoter DNA to which NM23-H2 binds is unorthodox and is essentially unknown, modelling with B-DNA may not be fruitful in this case. The structure of a cocrystal of NM23-H2 with the c-*myc* promoter sequence should clarify these interactions and any others not detected by this genetic study.

Chemical crosslinking data support the present model by suggesting that at least two subunits of NM23-H2, i.e., a dimer, may be required for making contacts with DNA (Fig. 4); although on the basis of our present data we cannot rule out the presence of higher order complexes on the DNA. Photocrosslinking of endogenous NM23-H2 in B-cell extracts (17) have also indicated that a 34-kDa species (dimer) is making contact with c-*myc* DNA, as well as the active component in the deregulation of the c-*myc* gene by NM23-H2 in Burkitt lymphoma cells.

The model for DNA binding by NM23-H2 presented here is inconsistant with theoretical models proposed earlier (3, 26), as well as being incompatible with a more recent model suggested by Webb *et al.* (29), in which DNA binding occurs on the side of the molecule covered by helices αA and α2 also involving the C-terminal fragment, and with that of Moréra *et al.* (28), in which DNA binds to a palm-like domain involving β sheets, helices αA, α2, and α4. These latter models imply the existence of six DNA-binding domains and/or require major conformational changes in the hexamer. The model we present here has an experimental basis as well as being simpler, in that DNA binding requires no major conformational changes and the functional subunit making contact with DNA is the stable dimer. Our model also implies that in addition to the physical separation of the catalytic and the DNA-binding domains, these two functions may be modulated by the oligomerization state of the protein: for example, DNA binding could be accomplished by the dimer, but for enzyme catalysis, the hexamer may be the active structure (30). If the DNA is presented with a hexamer but presumably only one of the dimers makes contact, this may explain the relatively high (μM) dissociation constant that we have observed in all of our DNA-binding studies.

In summary, we investigated the DNA-binding mechanism of NM23-H2/NDPK. Our results, based on site-directed mutational analysis, suggest that the DNA-binding specificity of NM23-H2 resides in nonconserved surface residues located on a different functional domain from that used for enzyme catalysis. In our model, NM23-H2 binds DNA as a dimer, with the DNA-binding region consisting of residues donated by two adjacent monomers. This provides a combinatorial DNA-binding surface located on the front "face" and along the 2-fold axis of the dimer (Fig. 5). Chemical crosslinking data support a dimeric DNA-binding mode.

We thank J. Janin for helpful discussions, for the modelling of R34 interactions, and for structural information including the coordinates for NM23-H2. We also thank I. Lascu for suggestions and for erythrocyte NDPKs and M. J. Lacombe for the NDPK-A cDNA clone. We are grateful to F. Hughson for making Fig. 5 and for discussions and S. J. Flint for review of the manuscript. The technical assistance of C. F. Conroy is acknowledged. This work was supported by National Institutes of Health Grant ROI CA5584 to E.H.P.

1. Steeg, P. S., Bevilacqua, G., Kopper, L., Thorgeirsson, U. P., Talmadge, J. E., Liotta, L. A. & Sobel, M. E. (1988) *J. Natl. Cancer Inst.* **80,** 199–206.
2. Steeg, P. S., De La Rosa, A., Flatow, U., MacDonald, N. J., Benedict, M. & Leone, A. (1993) *Breast Cancer Res. Treatment* **25,** 175–187.
3. Stahl, J. A., Leone, A., Rosengard, A. M., Porter, L., King, C. R. & Steeg, P. S. (1991) *Cancer Res.* **52,** 445–449.
4. Nakayama, T., Ohtsuru, A., Nakao, K., Shima, M., Nakata, K., Watanabe, K., Ishii, N., Kumura, N. & Nagataki, S. (1992) *J. Natl. Cancer Inst.* **84,** 1349–1354.
5. Nakayama, H., Yasui, W., Yokozaki, H. & Tahara, E. (1993) *Jpn. J. Cancer Res.* **84,** 184–190.
6. Hailat, N., Keim, D. R., Melhem, R. F., Zhu, X., Eckerskorn, C., Broudeur, G. M., Reynolds, C. P., Seeger, R. C., Lottspeich, F., Strahler, J. R. & Hanash, S. M. (1991) *J. Clin. Invest.* **88,** 342–345.
7. Leone, A., Seeger, R. C., Hong, C. M., Hu, Y. Y., Arboleda, M. J., Broderu, G. M., Stram, D., Slamon, D. J. & Steeg, P. S. (1993) *Oncogene* **8,** 855–865.
8. Keim, D., Hailat, N., Melhem, R., Zhu, X. X., Lascu, I., Veron, M., Strahler, J. & Hanash, S. M. (1992) *J. Clin. Invest.* **89,** 919–924.
9. Ohneda, K., Fukuda, M., Shimada, N., Ishikawa, N., Ichou, T., Kaji, K., Toyota, T. & Kimura, N. (1994) *FEBS Lett.* **348,** 273–277.
10. Dearolf, C. R., Tripoulas, N., Biggs, J. & Shearn, A. (1988) *Dev. Biol.* **129,** 169–178.
11. Biggs, J., Tripoulas, N., Hersperger, E., Dearolf, C. & Sheam, A. (1988) *Genes Dev.* **2,** 1333–1343.
12. Okabe-Kado, J., Kasukabe, T., Honma, Y., Hayashi, M., Henzel, W. J. & Hozumi, M. (1992) *Biochem. Biophys. Res. Commun.* **182,** 987–994.
13. Okabe-Kado, J., Kasukabe, T., Hozumi, M., Honma, Y., Kimura, N., Baba, H., Urano, T. & Hiroshi, S. (1995) *FEBS Lett.* **363,** 311–315.
14. Venturelli, D., Mártinez, R., Melotti, P., Casella, I., Peschle, C., Cucco, C., Spampinato, G., Darzynkiewicz, Z. & Calabretta, B. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7435–7439.
15. Postel, E. H., Berberich, S. J., Flint, S. J. & Ferrone, C. A. (1993) *Science* **261,** 478–480.
16. Berberich, S. J. & Postel, E. H. (1995) *Oncogene* **10,** 2343–2347.
17. Ji, L., Arcinas, M. & Boxer, L. M. (1995) *J. Biol. Chem.* **270,** 13392–13398.
18. Agarwal, R. P., Robinson, B. & Parks. R. E. (1978) *Methods Enzymol.* **51,** 376–386.
19. Postel, E. H. (1996) *Attempts to Understand Metastasis Formation:* Vol II, Regulatory Factors, Current Topics In Microbiology and Immunology, in press.
20. Backer, J. K., Mendola, C. E., Kovesdi, I., Fairhurst, J. L., O'Hara, B. & Eddy, R. L., Jr. (1993) *Oncogene* **8,** 497–502.
21. Howlett, A. R., Petersen, O. W., Steeg, P. S. & Bissell, M. J. (1994) *J. Natl. Cancer Inst.* **86,** 1838–1844.
22. Gilles, A.-M., Presecan, E., Vonica, A. & Lascu, I. (1991) *J. Biol. Chem.* **266,** 8784–8789.
23. Arcinas, M. & Boxer, L. M. (1994) *Oncogene* **9,** 2699–2706.
24. Postel, E. H., Mango, S. E. & Flint, S. J (1989 *Mol. Cell. Biol.* **9,** 5123–5133.
25. Dumas, C., Lascu, I., Moréra, S., Glaser, P., Fourme, R., Wallet, V., Lacombe, M.-L., Véron, M. & Janin, J. (1992) *EMBO J.* **11,** 3203–3208.
26. Chiadmi, M., Moréra, S., Lascu, I., Dumas, C., Le Bras, G., Véron, M. & Janin, J. (1993) *Structure* **1,** 283–293.
27. Moréra, S., LeBras, G., Lascu, I., Lacombe, M. L., Véron, M. & Janin, J. (1994) *J. Mol. Biol.* **243,** 873–890.
28. Moréra, S., Lacombe, M.-L., Yingwu, X., LeBras, G. & Janin, J. (1995) *Structure* **3,** 1307–1314.
29. Webb, P. A., Perisic, O., Mendola, C. E., Backer, J. M. & Williams, R. L. (1995) *J. Mol. Biol.* **251,** 574–587.
30. Lascu, I., Chaffotte, A., Limbourg-Bouchon, B. & Véron, M. (1992) *J. Biol. Chem.* **268,** 20268–20275.
31. Sastre-Garau, X., Lacombe, M. L., Jouve, M., Véron, M. & Magdelénat, J. (1992) *Int. J. Cancer* **50,** 533–538.
32. Spencer, C. A. & Groudine, M. (1991) *Adv. Cancer Res.* **56,** 1–48.
33. Marcu, K. B., Bossone, S. A. & Patel, A. J. (1992) *Annu. Rev. Biochem.* **61,** 809–860.
34. Prendergast, G. C. & Ziff, E. B. (1992) *Trends Genet.* **8,** 91–97.
35. Postel, E. H. & Ferrone, C. A. (1994) *J. Biol. Chem.* **269,** 8627–8630.
36. Hildebrandt, M., Lacombe, M.-L., Passeron, S. & Véron, M. (1995) *Nucleic Acids Res.* **23,** 3858–3864.
37. Suzuki, M. & Yagi, N. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 12357–12361.
38. Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. (1994) *Science* **265,** 346–355.
39. McDonald, N. J., De La Rosa, A., Benedict, M. A., Freije, J. M. P., Kritch, H. & Steeg, P. S. (1993) *J. Biol. Chem.* **268,** 25780–25789.