JBC RESEARCH ARTICLE

Structural basis of the TAM domain of BAZ2A in binding to DNA or RNA independent of methylation status

Received for publication, August 2, 2021, and in revised form, October 22, 2021 Published, Papers in Press, October 27, 2021,
https://doi.org/10.1016/j.jbc.2021.101351

Sizhuo Chen[1,‡], **Mengqi Zhou**[1,‡], **Aiping Dong**[2], **Peter Loppnau**[2], **Min Wang**[3], **Jinrong Min**[1,2,4,*], **and Ke Liu**[1,*]

From the [1]*Hubei Key Laboratory of Genetic Regulation and Integrative Biology, School of Life Sciences, Central China Normal University, Wuhan, PR China;* [2]*Structural Genomics Consortium, University of Toronto, Toronto, Ontario, Canada;* [3]*Testing & Analysis Center, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, PR China;* [4]*Department of Physiology, University of Toronto, Toronto, Ontario, Canada*

Edited by Patrick Sung

Bromodomain adjacent to zinc finger domain protein 2A (BAZ2A) (also called transcription termination factor-1 interacting protein 5), a key component of the nucleolar remodeling complex, recruits the nucleolar remodeling complex to ribosomal RNA genes, leading to their transcriptional repression. In addition to its tandem plant homeodomain–bromodomain that is involved in binding to acetylated histone H4, BAZ2A also contains a methyl-CpG–binding domain (MBD)-like Tip5/ARBP/MBD (TAM) domain that shares sequence homology with the MBD. In contrast with the methyl-CpG–binding ability of the canonical MBD, the BAZ2A TAM domain has been shown to bind to promoter-associated RNAs of ribosomal RNA genes and promoter DNAs of other genes independent of DNA methylation. Nevertheless, how the TAM domain binds to RNA/DNA mechanistically remains elusive. Here, we characterized the DNA-/RNA-binding basis of the BAZ2A TAM domain by EMSAs, isothermal titration calorimetry binding assays, mutagenesis analysis, and X-ray crystallography. Our results showed that the TAM domain of BAZ2A selectively binds to dsDNA and dsRNA and that it binds to the backbone of dsDNA in a sequence nonspecific manner, which is distinct from the base-specific binding of the canonical MBD. Thus, our results explain why the TAM domain of BAZ2A does not specifically bind to mCG or TG dsDNA like the canonical MBD and also provide insights for further biological study of BAZ2A acting as a transcription factor in the future.

Bromodomain adjacent to zinc finger domain protein 2A (BAZ2A) (or transcription termination factor-1 interacting protein 5) belongs to the bromodomain adjacent to zinc finger proteins (BAZ) family of chromatin remodeling factors, which function in chromatin remodeling, DNA replication, and DNA repair (1–3). BAZ2A was first identified as a partner of transcription termination factor 1, a transcription factor regulating transcription of RNA polymerase I–mediated ribosomal RNA genes (rDNAs) (4). As the largest component of the nucleolar remodeling complex (NoRC), BAZ2A targets the NoRC to rDNAs by binding to promoter-associated RNAs (pRNAs) and acetylated H4K16 and represses transcription of rDNAs by establishing repressive heterochromatic structures at rDNA promoters (5–10). BAZ2A can also regulate expression of other genes by recruiting the NoRC-mediated DNA methylation and EZH2-mediated histone H3K27me3 methylation activities to the target loci (11, 12). Overexpression of BAZ2A has been reported in patients with prostate cancer and contributes to poor prognosis and cancer recurrence (11, 13). BAZ2A is also upregulated in hepatocellular carcinoma (14) and chronic lymphocytic leukemia (15), and an FSIP1–BAZ2A fusion gene has been identified in multiple endocrine neoplasia type 2A (16).

BAZ2A contains multiple functional domains, including DNA-binding homeobox and different transcription factors, WHIM (WSTF, HB1, Itc1p, MBD9), plant homeodomain–bromodomain PHD-BRD, and Tip5/ARBP/MBD (TAM) domains (Fig. 1*A*). Most of these functional domains have also been found in other chromatin-modifying complexes, such as ATP-utilizing chromatin assembly and remodeling factor complex, chromatin-accessibility complex, and WSTF-ISWI chromatin remodeling complex (17–22), suggesting that they might share similar functions. The tandem PHD finger and BRD domain of BAZ2A is the most extensively studied domain. The PHD-BRD tandem domain of BAZ2A is essential for rDNA silencing with its abilities in histone H3/H4 binding and chromatin-modifying enzyme recruitment (10, 23). The TAM domain of BAZ2A is a methyl-CpG–binding domain (MBD)-like domain. In comparison with the canonical MBD, the TAM domain contains an additional C-terminal extension of two β-strands and one α-helix (Fig. 1, *A* and *B*). Methyl-CpG binding protein 2 is the first identified MBD-containing protein, and its MBD was first shown to bind to unmethylated matrix/scaffold attachment regions through a consensus sequence of 5′-GGTGT-3′ and was also shown at almost the same time to recognize methylated CpG DNA (24–26). This domain is also found in other MBD-containing proteins, including MBD1-6 and histone-lysine N-methyltransferase SETDB1/2. However, different from specific mCG and TG DNA binding of the canonical MBD, such as those of MBD2 and methyl-CpG binding protein 2, previous data have suggested that the TAM domain of BAZ2A binds to the stem–loop structure of pRNA to mediate the chromatin localization of

‡ These authors contributed equally to this work.
* For correspondence: Ke Liu, keliu2015@mail.ccnu.edu.cn; Jinrong Min, jr.min@utoronto.ca.

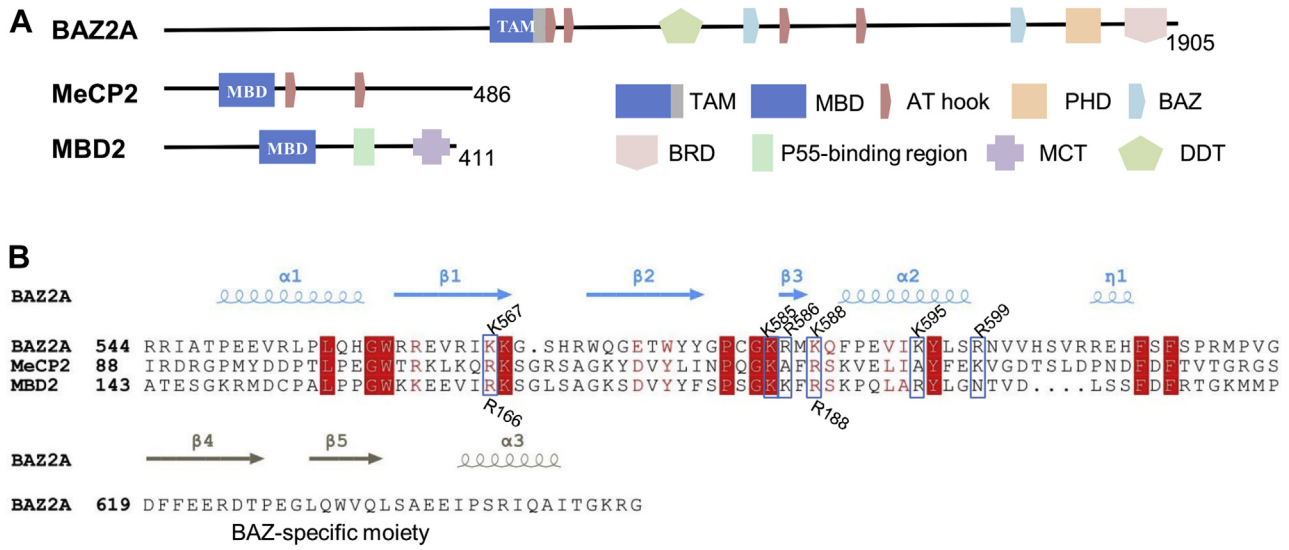

Figure 1. Sequence alignment of human BAZ2A TAM domain with MBDs of human MeCP2 and MBD2. *A*, domain organizations of human BAZ2A (NP_001191065.1), MeCP2 (NG_007107.2), and MBD2 (NP_003918.1). *B*, sequence alignment of human BAZ2A TAM domain with MBDs of human MeCP2 and MBD2. Secondary structure elements and dsDNA-binding residues of the BAZ2A TAM domain are indicated at the top of its sequence. The DNA-binding residues from MBD2 MBD are labeled at the bottom of the sequence. The BAZ-specific moiety of BAZ2A TAM domain is colored in *gray*. BAZ, bromodomain adjacent to the zinc finger proteins; BRD, bromodomain; DDT, DNA-binding homeobox and different transcription factors; MBD, methyl-CpG–binding domain; MCT, the C-terminal domain of methyl-CpG–binding protein 2 and 3; MeCP2, methyl-CpG binding protein 2; PHD, plant homeodomain finger; TAM, Tip5/ARBP/MBD.

the NoRC (9, 27) and also binds to DNA weakly in a methylation-independent manner (4). Considering that the TAM domain of BAZ2A is required and sufficient to mediate the nuclear matrix DNA binding (28) and mutations of the pRNA-binding residues, such as W531G and Y532A in mouse BAZ2A, do not influence the DNA-binding ability of mouse BAZ2A (9), it is plausible that the TAM domain of BAZ2A bound to DNA in a mode different than that of pRNA (9).

To investigate the DNA-/RNA-binding ability of the TAM domain of BAZ2A and its structural basis, we carried out *in vitro* DNA-/RNA-binding assays and determined crystal structures of the BAZ2A TAM domain in an apo-state and complex with two 12mer dsDNAs. Our binding results showed that the TAM domain of BAZ2A selectively binds to dsDNA and dsRNA. Our structural data demonstrated that the TAM domain of BAZ2A binds to the backbone phosphate groups of dsDNA with no sequence selectivity, which sheds light on the molecular basis for the recruitment of chromatin by the BAZ2A TAM domain.

Results

The TAM domain of BAZ2A binds to dsDNA without sequence selectivity

The DNA-binding selectivity of canonical MBDs, including those of methyl-CpG binding protein 2 and MBD1-4, has been investigated by isothermal titration calorimetry (ITC) binding assays and structural analysis systematically by us (29–31) and the others (32–35). The TAM domain of BAZ2A shares sequence homology with the MBD, and its NMR structure reveals that the major difference between the TAM domain and MBD is that the TAM domain contains an extra small α/β motif (named as BAZ-specific moiety here) at its C terminus (27). NMR titration experiments show that the C-terminal BAZ-specific moiety of the BAZ2A TAM domain interacts with the double-stranded region of the stem–loop structure of pRNA in a sequence nonspecific manner (27). EMSAs in another study show that the TAM domain of BAZ2A binds weakly to dsDNA in a methylation-independent manner (4). To further investigate the binding selectivity of the BAZ2A TAM domain toward DNA and RNA, we carried out ITC and EMSA binding assays.

Our ITC and EMSA results showed that the TAM domain of BAZ2A bound to dsDNA and dsRNA, but not to ssDNA and ssRNA (Figs. 2 and S1), consistent with the report that the TAM domain of BAZ2A binds to the double-stranded hairpin structure of pRNA (27). Our binding results also showed that, unlike the canonical MBDs, BAZ2A binding to dsDNA is not dependent on cytosine methylation (Figs. 2*A* and S1). However, we found that the binding affinity for BAZ2A to dsRNA and dsDNA is weaker than that of pRNA previously measured by a filter binding assay (5 μM *via* 5 nM) (27). Because the filter binding assay was carried out using a longer construct including both the TAM domain and the following AT-hook and the latter is a putative DNA-/RNA- binding motif (27, 36), we proposed that the inclusion of the AT-hook might increase the binding affinity of BAZ2A to pRNA as well as the dsRNA and dsDNA.

Crystal structure of the TAM domain of BAZ2A in complex with dsDNA

To illustrate the binding mechanism of the TAM domain of BAZ2A to dsDNA or dsRNA, we tried to cocrystallize the TAM domain protein of BAZ2A with dsDNA or dsRNA and determined the crystal structures of the TAM domain of BAZ2A in its apo-form as well as in complex with two different 12mer

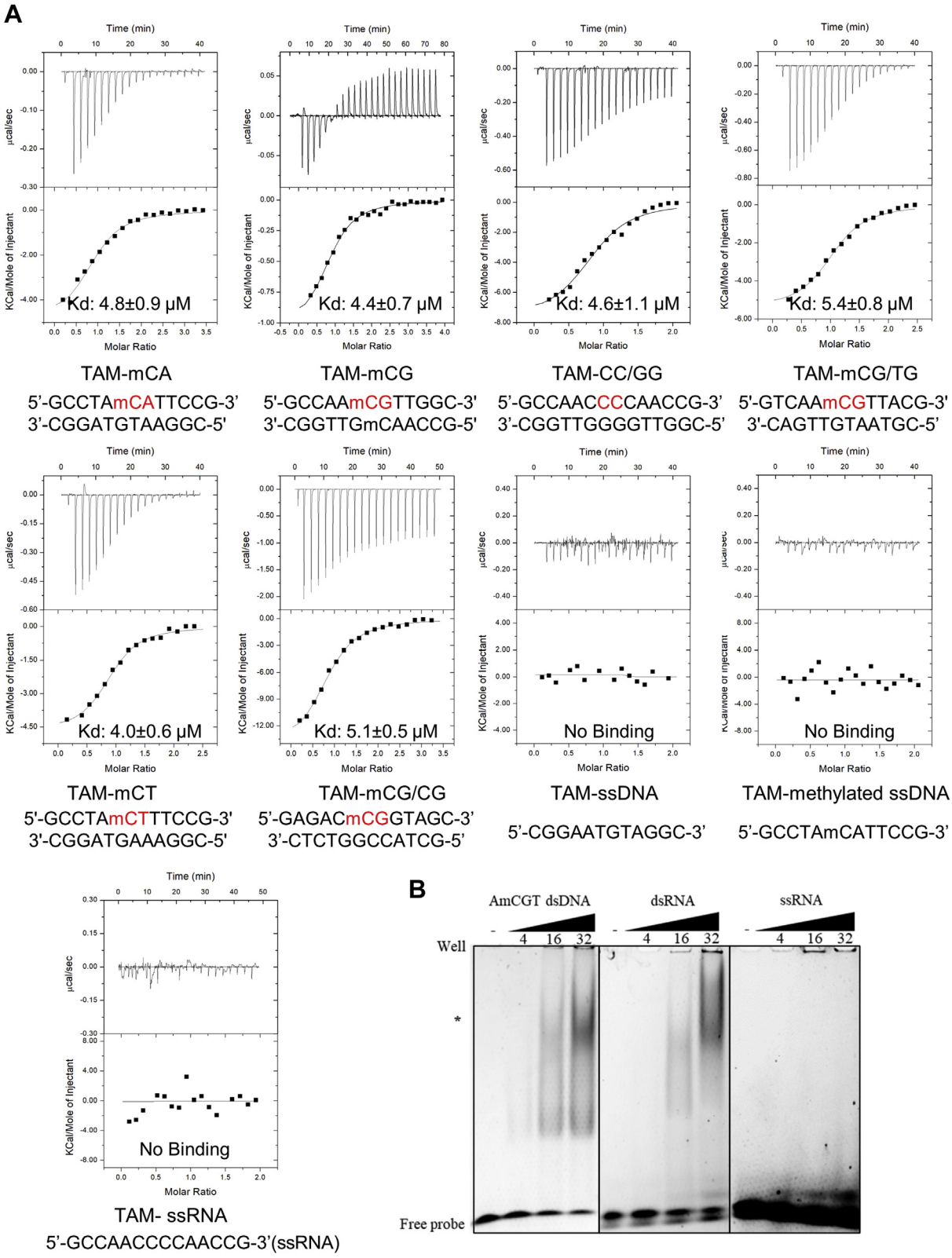

Figure 2. The TAM domain of BAZ2A binds to dsDNA and dsRNA without sequence selectivity. ***A,*** ITC binding curves for the TAM domain of BAZ2A to dsDNA, ssDNA, and ssRNA. The nucleotide sequences are shown below the ITC binding curves. *B*, the EMSAs of the BAZ2A TAM domain binding to DNA and RNA. In the EMSA, the concentration of DNA and RNA is 10 µM, and the concentrations of the protein in each set of four lanes are 0, 4, 16, and 32 µM, respectively. The DNA-/RNA-bound complex is highlighted by a *black asterisk*. ITC, isothermal titration calorimetry; TAM, Tip5/ARBP/MBD.

dsDNAs (named as mCG and mCA DNA here, Table 1). Consistent with the apo NMR structure, the apo crystal structure of the TAM domain of BAZ2A contains a twisted β-sheet comprised of five antiparallel β-strands, packing against four α-helices on one side of the β-sheet, which forms an MBD-like moiety and an additional BAZ-specific moiety (Figs. 3*A* and S2*A*). Two TAM molecules are found to bind antiparallelly to a single dsDNA duplex in both complex structures (Fig. 3, *B* and *C*). Superposition of the apo TAM structure with that of the TAM–dsDNA complexes reveals an RMSD value of approximately 0.35 Å over all the Cα atoms aligned, suggesting that the interactions with dsDNA do not lead to significant conformational changes of the TAM domain.

The canonical MBD binds to methylated CG or non-methylated TG DNA through both electrostatic interactions and arginine finger–mediated base-specific protein–DNA interactions using its central β-sheet (29–35). In contrast, in the BAZ2A TAM–dsDNA complex structures, the two TAM molecules provide a continuous positively charged surface to bind the backbone of dsDNA (Fig. 3, *D* and *E*). Our gel-filtration experiment showed that, when we increased the NaCl concentration of the buffer from 150 mM to 450 mM (Figs. 3*F* and S2*B*), the TAM–dsDNA complexes were dissembled, indicating that the salt-sensitive electrostatic interactions are essential for the BAZ2A–dsDNA complex formation.

Structural comparison of the BAZ2A–dsDNA complex with that of the MBD2–mCG complex reveals that the TAM domain of BAZ2A mainly interacts with dsDNA *via* β3 and α2 from its MBD-like moiety, which is different from the DNA-binding mode in the MBD2–mCG complex where MBD2 uses β2, β3, the loop linking β1 and β2, and α1 to recognize dsDNA (Fig. 3*G*). The corresponding MBD2–mCG DNA-binding region in the TAM domain is neutrally or even negatively charged, which explains why dsDNA could not bind to the TAM domain in a similar binding mode as observed in the MBD2–mCG complex (Fig. 3*H*). The previous NMR titrations have shown that the TAM domain of BAZ2A binds to pRNA mainly through β4, β5, and α3 from the BAZ-specific moiety (27), which is different from the dsDNA binding in our TAM–dsDNA structures, that is, the dsDNA binding region is far away from the BAZ-specific moiety in our complex structures (Fig. 3, *G* and *H*). We then generated a complex model of the TAM domain of BAZ2A bound to dsRNA and found that the dsRNA could be recognized using the same positively charged surface by the TAM domain (Fig. S2*C*). Thus, dsRNA might bind to the TAM domain of BAZ2A in a similar mode to that of dsDNA.

One BAZ2A TAM molecule is sufficient to bind to dsDNA

Because two TAM molecules are found to bind the same dsDNA duplex in the BAZ2A–dsDNA complex structure, we asked whether the dsDNA binding requires two BAZ2A TAM molecules. We first carried out gel-filtration chromatography analysis, which showed that the TAM protein alone behaved as a monomer in solution and the mixture of the TAM domain protein and dsDNA eluted at a volume corresponding to a molecular weight of ∼25 kDa (Figs. 3*F* and S2*B*). Then, we quantified the molecular weight of the TAM–dsDNA complex by multiangle light scattering coupled with size-exclusion chromatography (SEC-MALS) and obtained a molecular weight around 21 kDa for the mixture of the TAM domain and dsDNA (Fig. S3*A*), which is close to a theoretical molecular weight of the sum of a BAZ2A TAM molecule (13.9 kDa) and a dsDNA molecule (7.5 kDa). This finding is also consistent with our ITC data that the N values are approximately 1 when fitted with one set of site model (Fig. S1). Further analysis of the TAM structures revealed that the BAZ2A TAM molecule

Table 1
Data collection and refinement statistics

Protein	BAZ2A (aa 536–653)	BAZ2A (aa 536–653)	BAZ2A (aa 536–653)
PDB ID	7MWI	7FHJ	7MWL
DNA sequence		5′-GCCTA**mCAT**TCCG-3′ 3′-CGGATGTAAGGC-5′	5′-GCCAA**mCG**TTGGC-3′
Crystallization buffer	2.0 M Sodium formate, 0.1 M Tris, pH 8.5	20% PEG3350, 0.2 M ammonium chloride	10% PEG3350, 0.2 M L-proline, 0.1 M Hepes pH 7.5
Data collection			
Space group	P 6_1 2 2	P 6_1 2 2	P 6_1 2 2
Cell dimensions			
a, b, c (Å)	48.13, 48.13, 239.14	72.08, 72.08, 259.10	73.20, 73.20, 262.07
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	47.77–1.75 (1.78–1.75)	39.87–2.28 (2.36–2.28)	37.44–1.84 (1.88–1.84)
Completeness (%)	100.00 (100.00)	100.00 (100.00)	99.9 (99.1)
Rsymmetry	0.064 (1.140)	0.089 (1.119)	0.049 (0.951)
I/sigma I	34.5 (3.6)	23.8 (3.4)	48.5 (3.8)
CC1/2	1.000 (0.939)	0.999 (0.891)	1.000 (0.872)
Redundancy	20.1 (20.9)	20.6 (21.3)	35.5 (20.0)
Refinement			
Resolution (Å)	41.45–1.80	36.07–2.28	36.63–1.84
Reflections used	15,299/805	16,945/940	35,244/1855
No. of atoms/B-factor (Å^2)	993/33.58	2279/42.90	2400/42.0
Protein	914/33.72	1768/41.56	1777/40.76
DNA		487/48.25	488/46.22
Water	68/39.47	24/32.63	135/42.59
R work/free	0.185/0.213	0.218/0.247	0.213/0.230
RMSD bonds (Å)/angles (°)	0.019/1.93	0.009/1.67	0.013/1.34

Values in parentheses are for the highest-resolution shell.

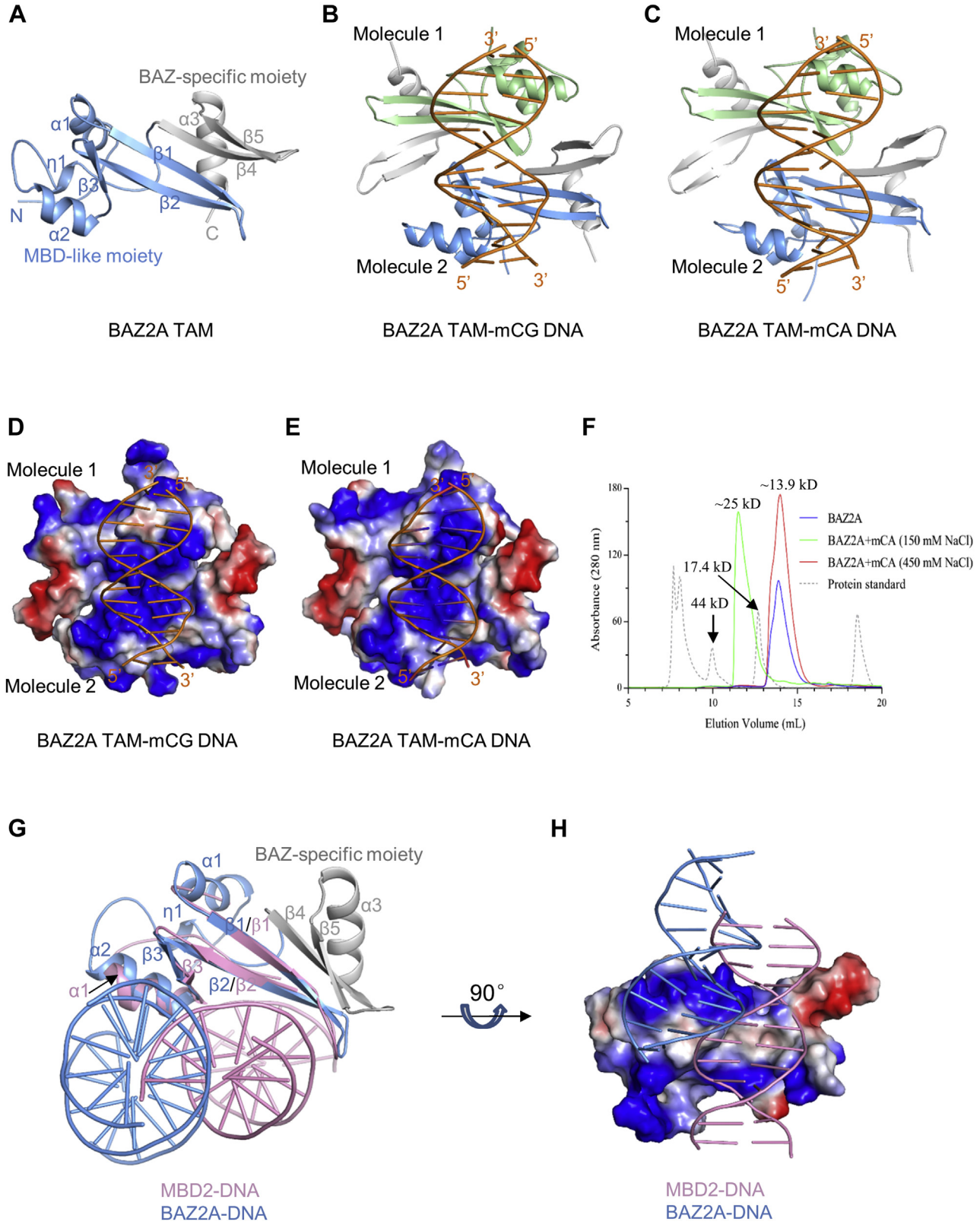

Figure 3. Crystal structures of the BAZ2A TAM domain in complex with dsDNA. *A*, the apo structure of the BAZ2A TAM domain in a cartoon representation. The BAZ-specific moiety and MBD-like moiety of the TAM domain are colored in *gray* and *blue*, respectively. *B* and *C*, overall structures of the BAZ2A TAM domain in complex with mCG and mCA dsDNA in a cartoon representation, respectively. The two BAZ2A TAM domain molecules are shown in *green* and *blue*, respectively. *D* and *E*, electrostatic surface representation of the BAZ2A TAM domain in complex with mCG and mCA dsDNA, respectively. F, size-exclusion chromatograms of the BAZ2A TAM domain in complex with mCA DNA in a buffer containing 150 mM or 450 mM NaCl. The molecular weights for the single TAM protein, TAM–dsDNA complex, and protein standards are labeled for each peak. G, superposition of the BAZ2A–mCG (*blue*) structure with that of the MBD2–mCG complex (*pink*, PDB code: 6CNP) in a cartoon representation. The BAZ-specific moiety of the TAM domain is colored in *gray*. H, superposition of the BAZ2A–mCG (*blue*) structure with that of the MBD2–mCG complex (*pink*, PDB code: 6CNP) in an electrostatic surface representation. MBD, methyl-CpG–binding domain; TAM, Tip5/ARBP/MBD.

packs with a symmetry-related TAM molecule in the apo crystal structure in a similar way to that of the two TAM molecules in the TAM–dsDNA complex structure, implying that the two TAM molecules observed in the complex are due to crystallization packing (Fig. S3*B*). Taken together, our data suggest that the TAM domain protein and dsDNA form a 1:1 complex in solution, and one BAZ2A TAM molecule is sufficient to bind to dsDNA.

Detailed interactions between the BAZ2A TAM domain and dsDNA

Detailed structural analysis showed that the TAM domain of BAZ2A mainly interacts with the backbone phosphate of dsDNA. The positively charged residues K585, R586, K588, K595, and R599 from β3 and α2 of each TAM domain simultaneously make electrostatic interactions with the dsDNA backbone, where the two TAM molecules are fixed at one side of the DNA duplex and form a dimeric complex (Figs. 3, *D* and *E* and 4). In addition to electrostatic interactions, numerous hydrogen bonds also contribute to the stabilization of the BAZ2A TAM–dsDNA complex (Fig. 4).

In the BAZ2A TAM–mCG DNA structure, the residues K585 and R586 in the first TAM molecule together with the residue R586 from the second TAM molecule form water-mediated hydrogen bonds with the backbone phosphates of G6′, mC7′, and G10, respectively (Fig. 4*A*). In addition, a glycerol molecule from the lysis or crystallization buffer interacts with the side chain of K588 in the first TAM molecule and forms water-mediated hydrogen bonds with the backbone phosphates of dsDNA (Fig. 4*A*). In the TAM–mCA structure, similar to the TAM–mCG complex, K585, R586, K588, K595, and R599 contribute to the electrostatic interactions with the backbone phosphates, but less water-mediated hydrogen bonds between the TAM domain and backbone phosphates of dsDNA were observed because of its lower resolution of around 2.3 Å (Fig. 4*B*). In addition, a weak hydrogen bond was found between the side chain of K585 from the second TAM molecule and backbone phosphates (Fig. 4*B*). Taken together, electrostatic interactions and water-mediated hydrogen-bonding interactions between the backbone of dsDNA and the BAZ2A TAM domain are the major driving forces for the BAZ2A TAM-dsDNA binding, which also explains why the BAZ2A TAM domain binds to dsDNA independent of the cytosine methylation and DNA sequences.

Structural basis of the TAM domain of BAZ2A in lacking the specific binding ability to mCG DNA like that of the canonical MBD

Previous studies have shown that two conserved arginine fingers of the canonical MBD specifically interact with the guanine base and simultaneously form cation–π interactions with the pyrimidine ring of methylated cytosine or thymine (29, 31, 33, 35, 37). Sequence alignment showed that the conserved arginine residues, that is, the R166 and R188 of MBD2, are replaced with the same positively charged lysine residues K567 and K588 in the TAM domain of BAZ2A (Fig. 1*B*). To investigate the importance of the conserved arginine fingers in DNA-binding specificity, we mutated the two arginine fingers of MBD2 to lysine and measured their DNA-binding affinities by ITC. The WT MBD of MBD2 binds to mCG DNA with a K_d value around 0.9 μM (30), whereas replacement of R166 or R188 with lysine in MBD2 reduces the binding affinity by about 30-fold and 8-fold, respectively (Fig. 5, *A* and *B*). The double point mutant with both R166 and R188 mutated to lysine does not show detectable binding to mCG DNA (Fig. 5*C*). On the other hand, when we mutated both K567 and K588 in the TAM domain of BAZ2A to arginine and carried out ITC binding assay, we found that the BAZ2A K567R/K588R mutant displays a similar methylated DNA-binding ability to that of BAZ2A WT (Figs. 2*A* and 5*D*), suggesting that the two K567R/K588R arginine residues of the BAZ2A mutant are not essential for the mCG DNA recognition as the canonical MBD does.

Remarkably, superposition of the BAZ2A TAM domain structure with that of MBD2–mCG complex revealed that the side chain of K567, corresponding to R166 in MBD2, does not point to the mCG base pair as observed in the canonical MBD structures and clashes with the backbone of DNA (Fig. 6). Further structural analysis revealed that the K567 harboring β1 tilts toward β4 of the BAZ-specific moiety and the two β strands (β4 and β5) from the BAZ-specific moiety integrate with the three β strands from the MBD-like moiety to form a 5-strand β-sheet (Fig. 6). This finding is consistent with the report that the MBD-like moiety and the BAZ-specific moiety of the TAM domain form a stable and integral structural fold (27) and also in agreement with our findings that the BAZ2A TAM protein with the BAZ-specific moiety truncated was not stable or soluble. The formation of the 5-strand β-sheet in the TAM domain restricts the side-chain orientation of K587 from binding methylated DNA as the arginine finger in the canonical MBD does. Taken together, our binding data and structural analysis explain why the MBD-like TAM domain of BAZ2A does not specifically bind to mCG or TG dsDNA like the canonical MBD.

Discussion

To date, the canonical MBD has been studied extensively. As an MBD-like domain, the DNA-/RNA-binding mechanism of the TAM domain is still unclear. In this study, we demonstrated by binding assays that the TAM domain of BAZ2A does not bind to ssRNA/ssDNA but binds to dsRNA/dsDNA in a sequence nonspecific mode. Our structural studies also revealed how the TAM domain binds to dsDNA in a sequence-independent and DNA methylation–independent manner. Very likely, the TAM domain also binds to dsRNA in a similar mode. Although it was proposed to regulate rDNA transcription *via* binding to pRNA, BAZ2A has been found to regulate gene expression of non-rRNA genes in embryonic stem cells as a transcription factor (12). Consistent with its DNA sequence nonspecific binding ability, BAZ2A is enriched over large chromatin regions that extends up to several hundred kilobases of DNA without high levels of DNA methylation or distinct

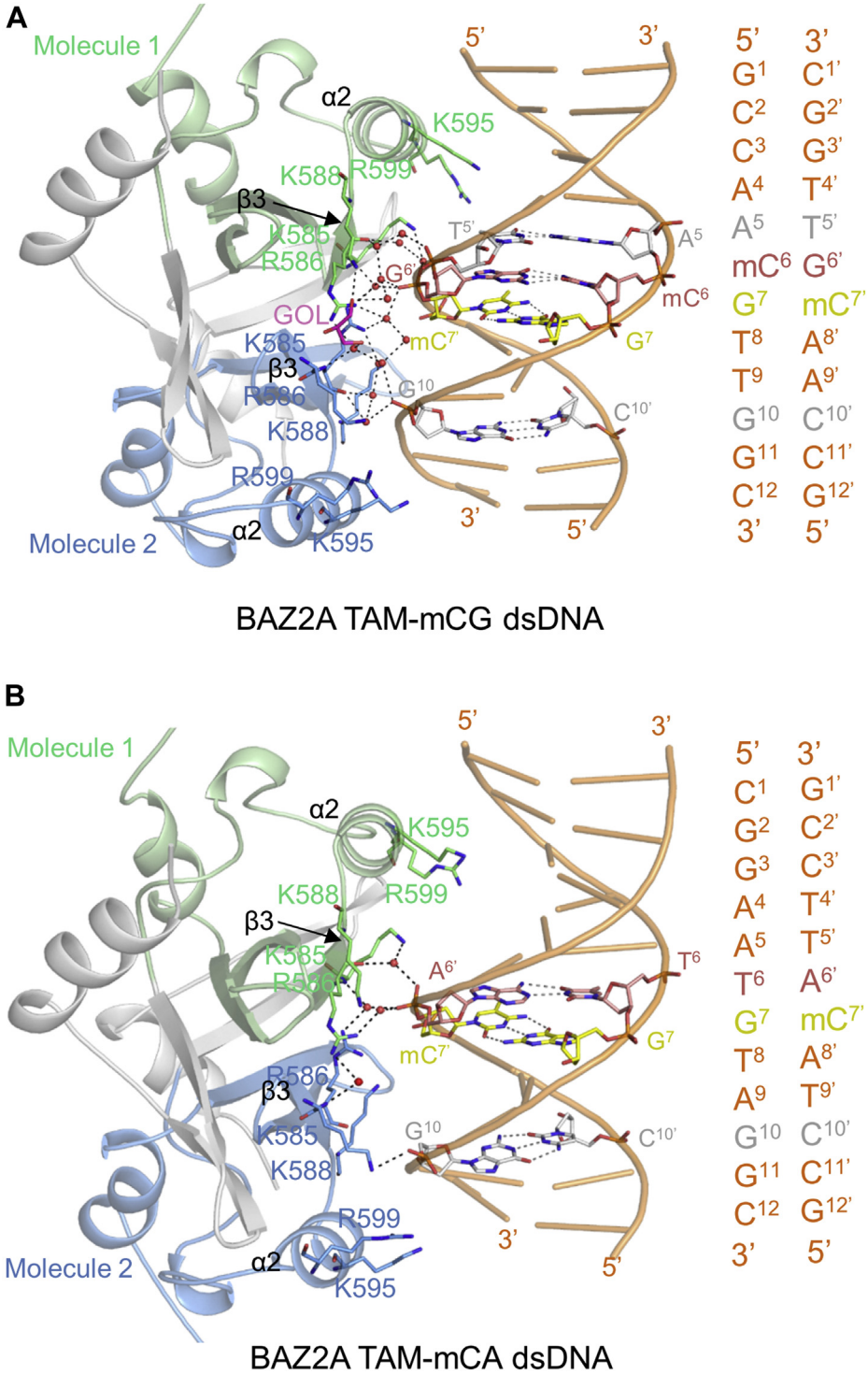

Figure 4. Detailed interactions between the BAZ2A TAM domain and dsDNA. *A*, overall structure of the BAZ2A TAM domain in complex with mCG DNA in a cartoon representation. The dsDNA-interacting residues are shown as stick models and colored in the same way as in Figure 3*B*. The DNA ligand is colored in *orange* except for the base pairs mC^{6}-$G^{6\prime}$ (*pink*), G^{7}-$mC^{7\prime}$ (*yellow*), A^{5}-$T^{5\prime}$ (*gray*), and G^{10}-$C^{10\prime}$ (*light gray*), which are shown as stick models. *B*, overall structure of the BAZ2A TAM domain in complex with mCA DNA in a cartoon representation. The dsDNA-interacting residues are shown as *stick* models and colored in the same way as in Figure 3*C*. The DNA ligand is colored in *orange* except for the base pairs T^{6}-$A^{6\prime}$ (*pink*), G^{7}-$mC^{7\prime}$ (*yellow*), and G^{10}-$C^{10\prime}$ (*light gray*). Hydrogen bonds formed between protein residues and dsDNA are marked as *black dashed lines*, whereas *gray dashed lines* represent hydrogen bonds between DNA base pairs. TAM, Tip5/ARBP/MBD.

binding peak profiles typical of transcription factors (12). Given that BAZ2A can recruit chromatin-modifying enzymes, such as HDAC1, DNMT1, DNMT3B, EZH2, as well as TOP2A and RNA helicase DHX9, to its target genes and regulate the chromatin conformation and epigenetic patterns at these loci, we speculated that the nonspecific binding of the TAM domain of BAZ2A to chromosomal DNA might play an important role in recruiting these enzymes (10, 12, 38).

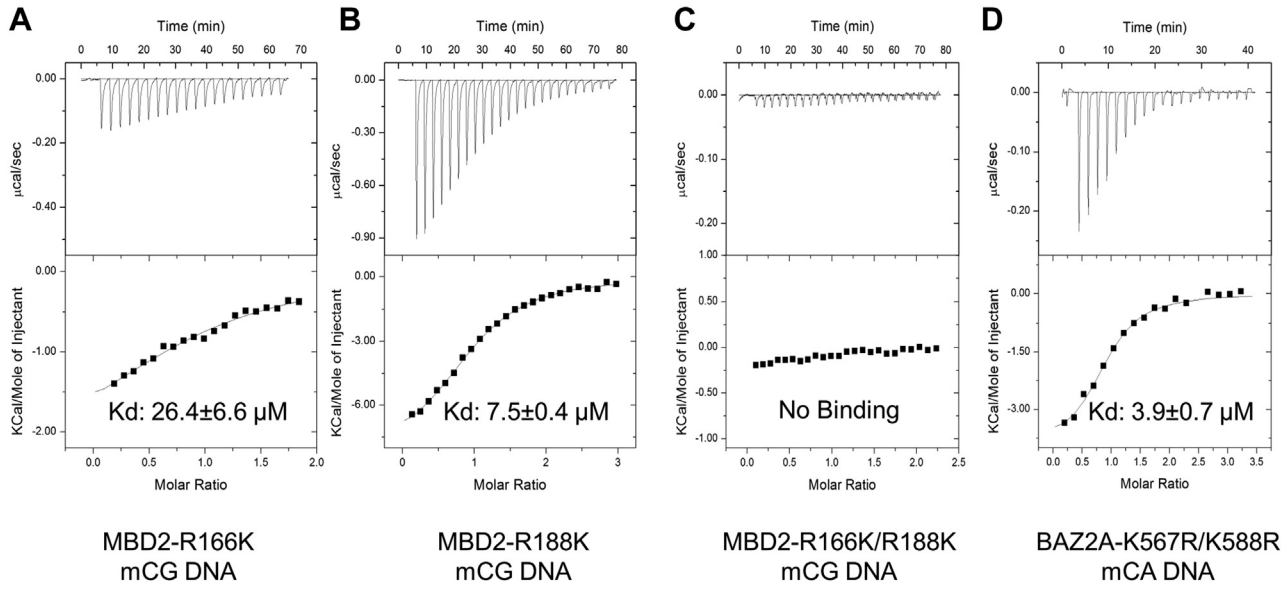

Figure 5. ITC binding results of MBD2 and BAZ2A mutants to methylated DNA. *A–C*, the ITC binding curves of the MBD2 MBD mutants with mCG DNA. The palindromic DNA sequence is 5′-GCCAAmCGTTGGC-3′. *D*, the ITC binding curve of the BAZ2A K567R/K588R mutant with mCA DNA. The palindromic DNA sequence is 5′-GCCTAmCATTCCG-3′/5′-CGGAATGTAGGC-3′. ITC, isothermal titration calorimetry; MBD, methyl-CpG–binding domain.

The sequence nonspecific protein–DNA interactions have also been found for other proteins, such as the protein involved in DNA replication and repair. For example, vaccinia virus uracil-DNA glycosylase (also called D4), acting as a DNA repair enzyme, can remove the uracil bases of DNA through hydrolyzing the N-glycosidic bond of deoxyuridine (39). In addition, D4 also serves as one component of the processive DNA polymerase complex in DNA replication, where A20 protein tethers D4 to DNA polymerase E9 (40). Although the glycosylase activity of D4 is not essential for the processive DNA synthesis (41), only short stretches of DNA are synthesized by polymerase E9 in the absence of D4 (40). It was reported that D4 slides on DNA and repairs the damaged bases during DNA replication (40). The structure of poxvirus D4 in complex with an undamaged dsDNA revealed that two D4 molecules bind simultaneously to a DNA double helix, and each D4 molecule binds to one DNA strand (Fig. S4) (42). Electrostatic potential distribution analysis showed that D4 provides a positively charged surface to accommodate the negatively charged phosphate backbone of the dsDNA, which together with the hydrogen bonds formed by the protein residues and backbone phosphates stabilize the D4–dsDNA complex (Fig. S4) (42). This nonspecific DNA-binding mode of D4 might facilitate the DNA sliding and base repair during DNA replication.

Experimental procedures

Protein expression and purification

Human BAZ2A (aa 536–653) and MBD2 (aa 143–220) were subcloned into the pET28-MHL vector to generate N-terminal His-tagged fusion proteins with a tobacco etch virus cleavage site. The BAZ2A and MBD2 mutants were obtained by QuikChange site-directed mutagenesis (Agilent Technologies) using the WT BAZ2A (aa 536–653) and MBD2 (aa 143–220) expression constructs as templates, respectively. The plasmids were transformed into *Escherichia coli* BL21 (DE3), and the cells were induced with 0.5 mM IPTG at 14 °C overnight. The harvested cells were resuspended in the buffer containing 20 mM Tris HCl, pH 7.5, 500 mM NaCl, and 5% glycerol, which was then sonicated to break the cells. The supernatant of the cell lysis was collected after centrifugation at 10,000 rpm and further purified using the Ni-NTA resin (Qiagen). Purified proteins were treated with Tobacco etch virus protease to remove the His-tag followed by affinity chromatography, anion-exchange chromatography, and gel-filtration column chromatography (GE Healthcare). Finally, the purified proteins were concentrated to ∼10 mg/ml in 20 mM Tris HCl, pH 7.5, and 150 mM NaCl, with or without 1 mM DTT.

ITC assays

The ITC assays were performed by MicroCal iTC200 (Malvern) at 25 °C. All the DNA oligos used in this study were synthesized by IDT (Integrated DNA Technologies). The DNA was dissolved in a buffer of 20 mM Tris HCl, pH 7.5, and 150 mM NaCl, which is identical to the protein buffer. After adjusting the pH of DNA solution to around 7.5 using NaOH, the ssDNAs were annealed into a DNA duplex as described before (43). The protein and DNA oligo concentrations for the ITC experiment range from 30 to 60 µM and 0.5 mM to 1 mM, respectively, which were determined using UV absorbance (ND-1000 spectrophotometer). The K_d values were determined using a one-site binding fitting model by Origin 7.0 (MicroCal Inc).

EMSA

The palindromic dsDNA oligo 5′-GCCAA(mC)GTTGGC-3′ and RNA oligo 5′-GCCAACCCCAACCG-3′ and 5′-CGGU UGGGGUUGGC-3′ were synthesized with fluorescent FAM and CY3 at their 5′ ends (Integrated DNA Technologies), respectively. The DNA and RNA oligos were dissolved in a buffer

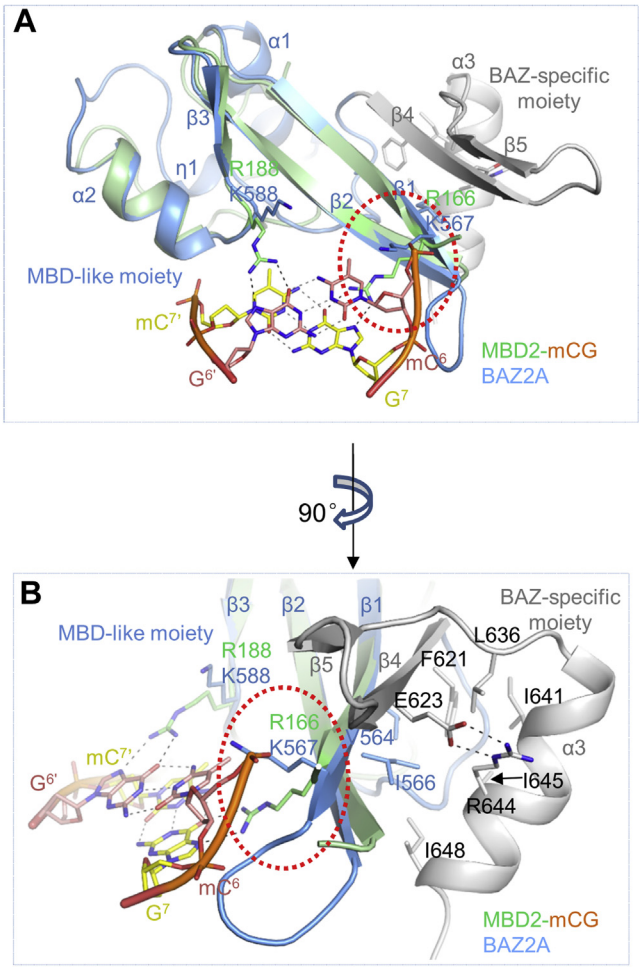

Figure 6. The MBD-like moiety and BAZ-specific moiety of the TAM domain form a stable and integral structure. *A*, structure comparison of the BAZ2A TAM domain (*blue*) and MBD2 MBD (*green*, PDB code: 6CNQ). The BAZ-specific moiety of the TAM domain is colored in *gray*. *B*, a different view of Figure 6*A* after 90° rotation. The DNA-interacting residues and protein-interacting DNA base pairs are shown as *stick* models. Hydrogen bonds formed between protein residues and DNA are marked as *black dashed lines*, whereas hydrogen bonds between base pairs are marked as *gray dashed lines*. BAZ, bromodomain adjacent to zinc finger proteins; MBD, methyl-CpG–binding domain; TAM, Tip5/ARBP/MBD.

containing 20 mM Tris HCl, pH 7.5, and 150 mM NaCl, with or without diethyl pyrocarbonate. They were then annealed to form duplexes by heating at 95 °C for 5 min and slow cooling to 16 °C. For EMSAs, proteins were incubated with FAM-labeled DNA or CY3-labeled RNA probe for 30 min at 4 °C. The samples were then analyzed on 6% native acrylamide gels in 0.5× TBE buffer under an electric field of 100 V for 90 min. Gels were visualized in the ethidium bromide mode on FluorChem R (Li-Cor).

SEC-MALS

We carried out the SEC shift assay to analyze the protein–DNA interactions. The BAZ2A (2 mg/ml) protein was mixed with dsDNA in a molar ratio of 1:1.1 and incubated on ice for 30 min. The samples and the protein molecular weight standard sample were analyzed by size-exclusion chromatograms (Superdex 75 10/300GL, GE) using the buffer containing 20 mM Tris HCl, pH 7.5, 1 mM DTT, and 150 mM or 450 mM NaCl. The SEC elution fractions corresponding to each peak were analyzed by SDS-PAGE.

For molar weight determination of the protein–DNA complex, the SEC-MALS experiment was performed using an HPLC-MALS system. A DAWN TREOS multiangle light scattering detector (Wyatt Technology) and an Optilab T-rEX refractometer (Wyatt Technology) were used in-line with WTC-015S5 hydrophilic film bonded silica column (Wyatt Technology) pre-equilibrated in a PBS buffer containing 50 mM Na_2HPO_4, 50 mM NaH_2PO_4, and 50 mM NaCl, pH 6.8, at a flow rate of 0.5 ml/min. Molecular weights of proteins were calculated with a dn/dc value (refractive index increment) of 0.185 ml/g using the Astra 6.1 program (Wyatt Technology).

Crystallization

For crystallization, 10 mg/ml BAZ2A TAM domain protein was used. To obtain the complex, BAZ2A TAM protein was mixed with dsDNA in a molar ratio of 1:1.1. The TAM domain protein of BAZ2A and the TAM–dsDNA complexes were crystallized using the sitting drop vapor diffusion method at 18 °C by mixing 0.5 μl of the samples with 0.5 μl of the reservoir solution. We obtained crystals for the apo-state of BAZ2A TAM domain, and its complexes with two 12mer dsDNAs, including the mCG DNA (5′-GCCAAmCGTTGGC-3′) and the mCA DNA (5′-GCCTAmCATTCCG-3′/5′-CGGAATG TAGGC-3′), respectively. The detailed crystallization conditions for each crystal are summarized in Table 1.

Data collection and structure determination

All the crystals were fast-frozen in liquid nitrogen with a cryoprotectant containing their respective reservoir solution with an additional 15%∼20% (v/v) glycerol. Diffraction images were collected at synchrotron or using rotating anode x-ray sources at 100 K and processed with the HKL suite (44) or XDS (45), and AIMLESS (46). Structures were solved by molecular replacement with PHASER (47) using coordinates from PDB entries 3C2I for the apo and complex structures of BAZ2A. Coot was used for model building (48), and model refinement was performed using PHENIX.REFINE (49) and REFMAC (50). The crystal data collection and structure refinement statistics are summarized in Table 1.

Data availability

Experimental data and coordinates for structures have been deposited in the Protein Data Bank with identifiers 7MWI, 7FHJ, and 7MWL.

Supporting information—This article contains supporting information.

Acknowledgments—We thank Dr W. Tempel for assistance in structure determination and H. Zhou of SSRF beamline BL17U1 for assistance in data collection.
This work was supported by funds from the National Natural Science Foundation of China (Number 31770834) and Central China

Normal University (CCNU) from the college's basic research and operation of the Ministry of Education (MOE) (Number CCNU19TS084). The Structural Genomics Consortium is a registered charity (no: 1097737) that receives funds from AbbVie, Bayer AG, Boehringer Ingelheim, Genentech, Genome Canada through Ontario Genomics Institute (OGI-196), the EU and EFPIA through the Innovative Medicines Initiative 2 Joint Undertaking (EUbOPEN Grant 875510), Janssen, Merck KGaA (aka EMD in Canada and the United States), Pfizer, Takeda, and the Wellcome Trust (106169/ZZ14/Z).

Author contributions—S. C., M. Z., P. L., M. W., and K. L. investigation; A. D. and K. L. formal analysis; K. L. and J. M. visualization; K. L. and J. M. supervision; K. L. and J. M. writing–original draft; S. C., M. Z., A. D. P. L., M. W., K. L. and J. M. writing–review and editing.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: BAZ, bromodomain adjacent to zinc finger proteins; BAZ2A, bromodomain adjacent to zinc finger domain protein 2A; ITC, isothermal titration calorimetry; MBD, methyl-CpG–binding domain; NoRC, nucleolar remodeling complex; PHD-BRD, plant homeodomain–bromodomain; pRNAs, promoter-associated RNAs; rDNAs, ribosomal RNA genes; SEC-MALS, size-exclusion chromatography; TAM, Tip5/ARBP/MBD..

References

1. He, X., F, H., Garlick, J. D., and Kingston, R. E. (2008) Diverse regulation of SNF2h chromatin remodeling by noncatalytic subunits. *Biochemistry* **47**, 7025–7033
2. Clapier, C. R., and Cairns, B. R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**, 273–304
3. Jones, M. H., Hamana, N., Nezu, J., and Shimane, M. (2000) A novel family of bromodomain genes. *Genomics* **63**, 40–45
4. Strohner, R., Nemeth, A., Jansa, P., Hofmann-Rohrer, U., Santoro, R., Langst, G., and Grummt, I. (2001) NoRC–a novel member of mammalian ISWI-containing chromatin remodeling machines. *EMBO J.* **20**, 4892–4900
5. Santoro, R., and Grummt, I. (2005) Epigenetic mechanism of rRNA gene silencing: Temporal order of NoRC-mediated histone modification, chromatin remodeling, and DNA methylation. *Mol. Cell Biol.* **25**, 2539–2546
6. Santoro, R., Li, J., and Grummt, I. (2002) The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. *Nat. Genet.* **32**, 393–396
7. Németh A, S. R., Grummt, I., and Längst, G. (2004) The chromatin remodeling complex NoRC and TTF-I cooperate in the regulation of the mammalian rRNA genes *in vivo*. *Nucleic Acids Res.* **32**, 4091–4099
8. Guetg, C., Scheifele, F., Rosenthal, F., Hottiger, M. O., and Santoro, R. (2012) Inheritance of silent rDNA chromatin is mediated by PARP1 via noncoding RNA. *Mol. Cell* **45**, 790–800
9. Mayer, C., Schmitz, K. M., Li, J., Grummt, I., and Santoro, R. (2006) Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol. Cell* **22**, 351–361
10. Zhou, Y., and Grummt, I. (2005) The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing. *Curr. Biol.* **15**, 1434–1438
11. Gu, L., Frommel, S. C., Oakes, C. C., Simon, R., Grupp, K., Gerig, C. Y., Bar, D., Robinson, M. D., Baer, C., Weiss, M., Gu, Z., Schapira, M., Kuner, R., Sultmann, H., Provenzano, M., *et al.* (2015) BAZ2A (TIP5) is involved in epigenetic alterations in prostate cancer and its overexpression predicts disease recurrence. *Nat. Genet.* **47**, 22–30
12. Dalcher, D., Tan, J. Y., Bersaglieri, C., Pena-Hernandez, R., Vollenweider, E., Zeyen, S., Schmid, M. W., Bianchi, V., Butz, S., Roganowicz, M., Kuzyakiv, R., Baubec, T., Marques, A. C., and Santoro, R. (2020) BAZ2A safeguards genome architecture of ground-state pluripotent stem cells. *EMBO J.* **39**, e105606
13. Pietrzak, K., Kuzyakiv, R., Simon, R., Bolis, M., Bar, D., Aprigliano, R., Theurillat, J. P., Sauter, G., and Santoro, R. (2020) TIP5 primes prostate luminal cells for the oncogenic transformation mediated by PTEN-loss. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 3637–3647
14. Li, C., Wu, W., Ding, H., Li, Q., and Xie, K. (2018) The transcription factor 7 like 2binding protein TIP5 activates betacatenin/transcription factor signaling in hepatocellular carcinoma. *Mol. Med. Rep.* **17**, 7645–7651
15. Hanlon, K., Rudin, C. E., and Harries, L. W. (2009) Investigating the targets of MIR-15a and MIR-16-1 in patients with chronic lymphocytic leukemia (CLL). *PLoS One* **4**, e7169
16. Du, Z. F., Li, P. F., Zhao, J. Q., Cao, Z. L., Li, F., Ma, J. M., and Qi, X. P. (2017) Genetic diagnosis of a Chinese multiple endocrine neoplasia type 2A family through whole genome sequencing. *J. Biosci.* **42**, 209–218
17. Eberharter, A., Vetter, I., Ferreira, R., and Becker, P. B. (2004) ACF1 improves the effectiveness of nucleosome mobilization by ISWI through PHD-histone contacts. *EMBO J.* **23**, 4029–4039
18. LeRoy, G., Orphanides, G., Lane, W. S., and Reinberg, D. (1998) Requirement of RSF and FACT for transcription of chromatin templates *in vitro*. *Science* **282**, 1900–1904
19. Ito, T., Levenstein, M. E., Fyodorov, D. V., Kutach, A. K., Kobayashi, R., and Kadonaga, J. T. (1999) ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. *Genes Dev.* **13**, 1529–1539
20. Bochar, D. A., Savard, J., Wang, W., Lafleur, D. W., Moore, P., Cote, J., and Shiekhattar, R. (2000) A family of chromatin remodeling factors related to Williams syndrome transcription factor. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 1038–1043
21. Poot, R. A., Dellaire, G., Hulsmann, B. B., Grimaldi, M. A., Corona, D. F., Becker, P. B., Bickmore, W. A., and Varga-Weisz, P. D. (2000) HuCHRAC, a human ISWI chromatin remodelling complex contains hACF1 and two novel histone-fold proteins. *EMBO J.* **19**, 3377–3387
22. Bozhenok, L., Wade, P. A., and Varga-Weisz, P. (2002) WSTF-ISWI chromatin remodeling complex targets heterochromatic replication foci. *EMBO J.* **21**, 2231–2241
23. Tallant, C., Valentini, E., Fedorov, O., Overvoorde, L., Ferguson, F. M., Filippakopoulos, P., Svergun, D. I., Knapp, S., and Ciulli, A. (2015) Molecular basis of histone tail recognition by human TIP5 PHD finger and bromodomain of the chromatin remodeling complex NoRC. *Structure* **23**, 80–92
24. von Kries, J. P., Buhrmester, H., and Stratling, W. H. (1991) A matrix/scaffold attachment region binding protein: Identification, purification, and mode of binding. *Cell* **64**, 123–135
25. Buhrmester, H., von Kries, J. P., and Stratling, W. H. (1995) Nuclear matrix protein ARBP recognizes a novel DNA sequence motif with high affinity. *Biochemistry* **34**, 4108–4117
26. Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69**, 905–914
27. Anosova, I., Melnik, S., Tripsianes, K., Kateb, F., Grummt, I., and Sattler, M. (2015) A novel RNA binding surface of the TAM domain of TIP5/BAZ2A mediates epigenetic regulation of rRNA genes. *Nucleic Acids Res.* **43**, 5208–5220
28. Zillner, K., Filarsky, M., Rachow, K., Weinberger, M., Langst, G., and Nemeth, A. (2013) Large-scale organization of ribosomal DNA chromatin is regulated by Tip5. *Nucleic Acids Res.* **41**, 5251–5262
29. Liu, K., Lei, M., Wu, Z., Gan, B., Cheng, H., Li, Y., and Min, J. (2019) Structural analyses reveal that MBD3 is a methylated CG binder. *FEBS J.* **286**, 3240–3254
30. Liu, K., Xu, C., Lei, M., Yang, A., Loppnau, P., Hughes, T. R., and Min, J. (2018) Structural basis for the ability of MBD domains to bind methyl-CG and TG sites in DNA. *J. Biol. Chem.* **293**, 7344–7354

31. Lei, M., Tempel, W., Chen, S., Liu, K., and Min, J. (2019) Plasticity at the DNA recognition site of the MeCP2 mCG-binding domain. *Biochim. Biophys. Acta Gene Regul. Mech.* **1862**, 194409
32. Ohki, I., Shimotake, N., Fujita, N., Jee, J., Ikegami, T., Nakao, M., and Shirakawa, M. (2001) Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell* **105**, 487–497
33. Ho, K. L., McNae, I. W., Schmiedeberg, L., Klose, R. J., Bird, A. P., and Walkinshaw, M. D. (2008) MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell* **29**, 525–531
34. Otani, J., Arita, K., Kato, T., Kinoshita, M., Kimura, H., Suetake, I., Tajima, S., Ariyoshi, M., and Shirakawa, M. (2013) Structural basis of the versatile DNA recognition ability of the methyl-CpG binding domain of methyl-CpG binding domain protein 4. *J. Biol. Chem.* **288**, 6351–6362
35. Scarsdale, J. N., Webb, H. D., Ginder, G. D., and Williams, D. C., Jr. (2011) Solution structure and dynamic analysis of chicken MBD2 methyl binding domain bound to a target-methylated DNA sequence. *Nucleic Acids Res.* **39**, 6741–6752
36. Filarsky, M., Zillner, K., Araya, I., Villar-Garea, A., Merkl, R., Langst, G., and Nemeth, A. (2015) The extended AT-hook is a novel RNA binding motif. *RNA Biol.* **12**, 864–876
37. Ren, R., Horton, J. R., Zhang, X., Blumenthal, R. M., and Cheng, X. (2018) Detecting and interpreting DNA methylation marks. *Curr. Opin. Struct. Biol.* **53**, 88–99
38. Leone, S., Bar, D., Slabber, C. F., Dalcher, D., and Santoro, R. (2017) The RNA helicase DHX9 establishes nucleolar heterochromatin, and this activity is required for embryonic stem cell differentiation. *EMBO Rep.* **18**, 1248–1262
39. Lindahl, T. (1974) An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 3649–3653
40. Boyle, K. A., Stanitsa, E. S., Greseth, M. D., Lindgren, J. K., and Traktman, P. (2011) Evaluation of the role of the vaccinia virus uracil DNA glycosylase and A20 proteins as intrinsic components of the DNA polymerase holoenzyme. *J. Biol. Chem.* **286**, 24702–24713
41. De Silva, F. S., and Moss, B. (2003) Vaccinia virus uracil DNA glycosylase has an essential role in DNA synthesis that is independent of its glycosylase activity: Catalytic site mutations reduce virulence but not virus replication in cultured cells. *J. Virol.* **77**, 159–166
42. Schormann, N., Banerjee, S., Ricciardi, R., and Chattopadhyay, D. (2015) Binding of undamaged double stranded DNA to vaccinia virus uracil-DNA Glycosylase. *BMC Struct. Biol.* **15**, 10
43. Xu, Y., Xu, C., Kato, A., Tempel, W., Abreu, J. G., Bian, C., Hu, Y., Hu, D., Zhao, B., Cerovina, T., Diao, J., Wu, F., He, H. H., Cui, Q., Clark, E., *et al.* (2012) Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. *Cell* **151**, 1200–1213
44. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326
45. Krug, M., Weiss, M. S., Heinemann, U., and Mueller, U. (2012) XDSAPP: A graphical user interface for the convenient processing of diffraction data usingXDS. *J. Appl. Crystallogr.* **45**, 568–572
46. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214
47. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
48. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
49. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367
50. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367