

Molecular basis of facilitated target search and sequence discrimination of TALE homeodomain transcription factor Meis1

Received: 22 February 2023

Accepted: 5 August 2024

Published online: 14 August 2024

 Check for updates

Seo-Ree Choi^{1,2,7}, Juyong Lee^{3,4,5,7}✉, Yeo-Jin Seo¹, Ho-Seong Jin¹, Hye-Bin Ahn¹, Youyeon Go¹, Nak-Kyoon Kim², Kyoung-Seok Ryu⁶ & Joon-Hwa Lee^{1,8}

Transcription factors specifically bind to their consensus sequence motifs and regulate transcription efficiency. Transcription factors are also able to non-specifically contact the phosphate backbone of DNA through electrostatic interaction. The homeodomain of Meis1 TALE human transcription factor (Meis1-HD) recognizes its target DNA sequences via two DNA contact regions, the L1- α 1 region and the α 3 helix (specific binding mode). This study demonstrates that the non-specific binding mode of Meis1-HD is the energetically favored process during DNA binding, achieved by the interaction of the L1- α 1 region with the phosphate backbone. An NMR dynamics study suggests that non-specific binding might set up an intermediate structure which can then rapidly and easily find the consensus region on a long section of genomic DNA in a facilitated binding process. Structural analysis using NMR and molecular dynamics shows that key structural distortions in the Meis1-HD–DNA complex are induced by various single nucleotide mutations in the consensus sequence, resulting in decreased DNA binding affinity. Collectively, our results elucidate the detailed molecular mechanism of how Meis1-HD recognizes single nucleotide mutations within its consensus sequence: (i) through the conformational features of the α 3 helix; and (ii) by the dynamic features (rigid or flexible) of the L1 loop and the α 3 helix. These findings enhance our understanding of how single nucleotide mutations in transcription factor consensus sequences lead to dysfunctional transcription and, ultimately, human disease.

Transcription factors (TFs) recognize consensus sequence motifs but are also able to interact with non-specific DNA through their positively charged side chains¹. Transcription efficiency depends on the ability of TFs to locate specific binding sites². Thus, it is very important for TFs to

accurately and quickly find their consensus binding site within an overwhelming amount of DNA. How do they do this? The specific binding of TFs is achieved by two types of intermolecular interactions: (i) sequence-independent electrostatic interaction with the phosphate

¹Department of Chemistry and RINS, Gyeongsang National University, Jinju, Gyeongsangnam-do 52828, Republic of Korea. ²Advanced Analysis Data Center, Korea Institute of Science and Technology, Seoul 02456, Republic of Korea. ³Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, and College of Pharmacy, Seoul National University, Seoul 08826, Republic of Korea. ⁴Research Institute of Pharmaceutical Science, College of Pharmacy, Seoul National University, Seoul 08826, Republic of Korea. ⁵Arontier Co., Seoul 06735, Republic of Korea. ⁶Protein Structure Research Team, Korea Basic Science Institute, Cheongju, Chungcheongbuk-do 28119, Republic of Korea. ⁷These authors contributed equally: Seo-Ree Choi, Juyong Lee. ⁸Deceased: Joon-Hwa Lee. ✉e-mail: nicole23@snu.ac.kr

backbone of DNA and (ii) direct contact with the bases at the consensus site^{1,3,4}. The structural features of these optimized interactions for the specific binding of various TFs have been proven in structural studies of TF–DNA complexes. In contrast to specific binding, non-specific binding is decidedly transient and is usually considered as a discontinuous transition among various electrostatically interacting states between TF and DNA¹. So far, the structural features of a single defined non-specifically bound state between TFs and DNA have been rarely reported.

The homeodomain (HD) is a highly conserved DNA-binding motif that widely exists in eukaryotic TFs^{2,5,6}. Myeloid ecotropic viral integration site-1 (Meis1) belongs to the three amino acid loop extension (TALE) HD family and consists of a DNA-binding TALE-HD toward the C-terminus and two protein-protein interaction domains toward the N-terminus (Fig. 1a)^{7–9}. The consensus binding sequence of Meis1, 5'-TGACAG-3' (hexameric motif) (Fig. 1b), was identified by the systematic evolution of ligands by exponential enrichment (SELEX) or random binding site selection^{10–12}. When Meis1 forms a heterodimer or

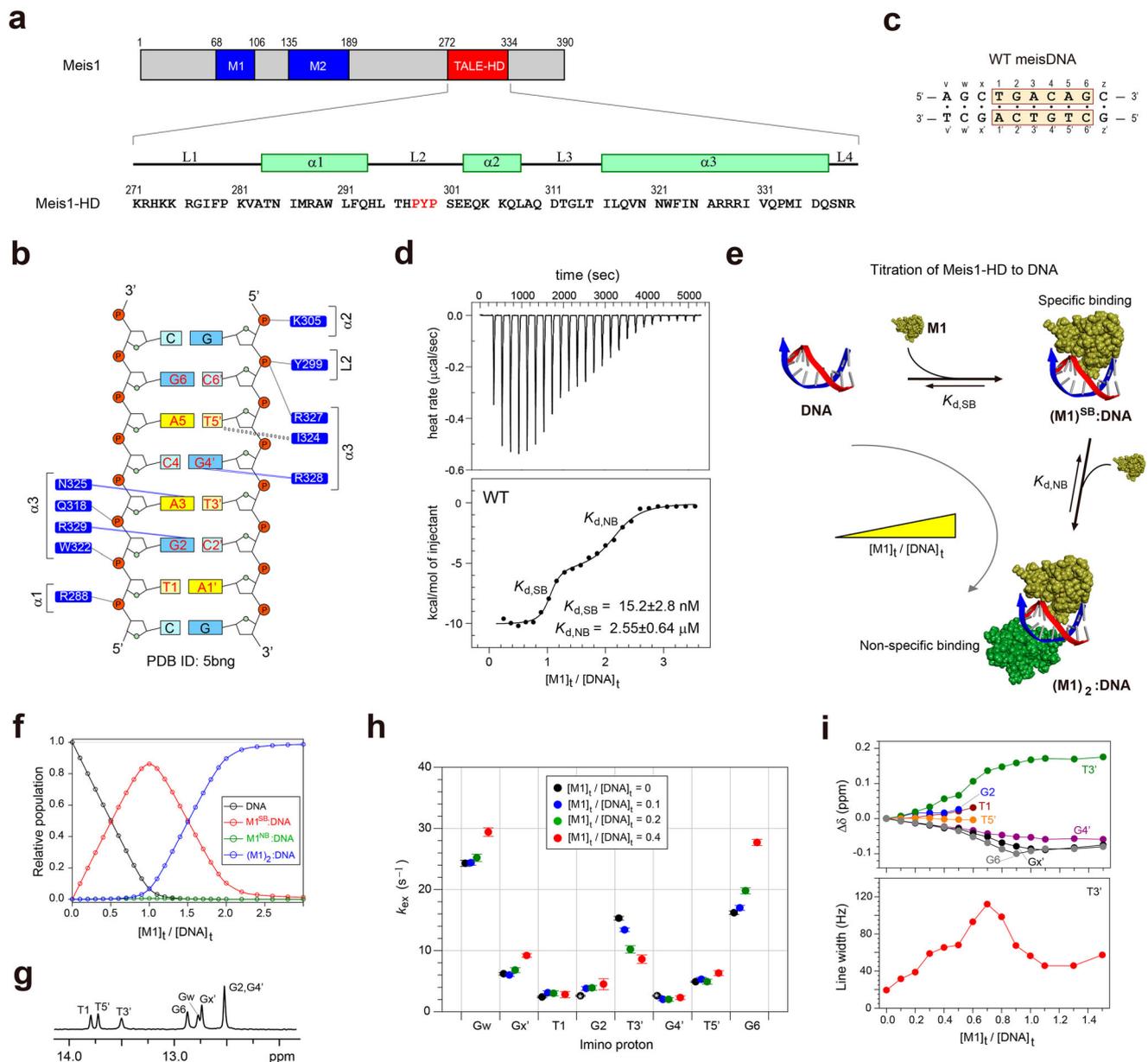


Fig. 1 | Dual DNA-binding mode of Meis1-HD. a Domain structure of human TF Meis1 and amino acid sequence of Meis1-HD. Numbering and secondary structure elements for Meis1-HD are shown on top of the sequence. The red color indicates TALE residues. **b** Residues of Meis1-HD involved in interaction with the consensus DNA duplex reported in a previous study¹². Intermolecular hydrogen bonds are indicated by solid lines and van der Waals contacts are shown by open circles. **c** Sequence context of WT meisDNA. Orange bars indicate consensus base-pairs for Meis1 binding. **d** The ITC binding isotherm of Meis1-HD titrated into WT meisDNA at 298 K. Raw heat data (top) and the integrated heat data with the nonlinear regression fit (bottom) are shown. **e** Binding modes of Meis1-HD (denoted as M1) to a 10-bp DNA duplex: (i) first, M1 specifically binds to the consensus region to form

an (M1)^{SB}:DNA complex; (ii) and then an (M1)₂:DNA complex is produced by non-specific binding of M1 to an (M1)^{SB}:DNA complex. **f** Relative populations of free WT meisDNA (black) and the Meis1-HD–DNA complexes (M1)^{SB}:DNA (red), (M1)^{NB}:DNA (green), or (M1)₂:DNA (blue) as a function of the $[M1]_t/[DNA]_t$ ratio. **g** 1D imino spectrum of WT meisDNA at 298 K. **h** The exchange rate constants of the imino protons of the WT meisDNA–Meis1-HD complexes at various $[M1]_t/[DNA]_t$ ratios. The error bars indicate the uncertainties associated with the curve fitting results. The asterisks indicate the k_{ex} value determined from the overlapped G2 and G4' imino proton resonances. **i** (Upper) ¹H chemical shift changes of the imino proton resonances and (lower) line width of the T3' imino proton resonance in the WT meisDNA complexed with Meis1-HD as a function of the $[M1]_t/[DNA]_t$ ratio.

Table 1 | Thermodynamic parameters and dissociation constants (K_d) for DNA binding of Meis1-HD at 298 K.^a

meisDNA	Specific binding				Non-specific binding			
	ΔH_{SB} (kcal/mol)	$-\Delta S_{SB}$ (kcal/mol)	n	$K_{d,SB}$ (nM)	ΔH_{NB} (kcal/mol)	$-\Delta S_{NB}$ (kcal/mol)	n	$K_{d,NB}$ (μ M)
WT	-10.18 ± 0.08	-0.49 ± 0.03	0.97 ± 0.01	15.2 ± 2.8	-5.60 ± 0.63	-2.04 ± 0.77	1.07 ± 0.03	2.55 ± 0.64
T1C	-7.00 ± 0.20	-3.11 ± 0.34	0.99 ± 0.03	39.8 ± 9.4	-5.56 ± 0.38	-2.16 ± 0.50	1.01 ± 0.07	2.22 ± 0.49
G2A	-8.60 ± 0.40	-0.89 ± 0.32	1.02 ± 0.03	112.9 ± 16.4	-5.06 ± 0.36	-2.36 ± 0.59	0.99 ± 0.03	3.75 ± 1.41
A3G	-5.89 ± 0.17	-3.81 ± 0.15	0.97 ± 0.03	78.1 ± 2.6	-5.29 ± 0.49	-2.32 ± 0.66	0.99 ± 0.01	2.68 ± 0.75
C4T	-8.35 ± 0.46	-1.27 ± 0.43	0.98 ± 0.01	90.2 ± 4.7	-5.40 ± 0.09	-2.38 ± 0.11	1.03 ± 0.05	1.96 ± 0.08
A5G	-7.46 ± 0.09	-2.51 ± 0.05	0.98 ± 0.01	49.7 ± 4.4	-4.97 ± 0.07	-2.76 ± 0.42	1.02 ± 0.07	2.36 ± 1.31

^aSample conditions: 10 mM sodium phosphate (pH 6.0), 100 mM NaCl.

heterotrimer with other TALE-HD or HD TFs, Meis1 displays distinctive DNA-binding specificity¹³. For example, Meis1–Pbx1 and Meis1–Drgx dimers preferentially bind the 5'-TGATTGACAG-3' and 5'-TAATTGACA-3' sequences, respectively^{11,12}. In contrast to in vitro activities, Meis1 displays low sequence specificity for DNA binding in vivo, with two major target sequences: the hexameric sequence identical to the in vitro consensus sequence and the 5'-WGATNNAT-3' octameric sequence similar to the consensus binding sites of the Pbx–Hox dimer, where W is T or A and N is any sequence¹³. In vivo, the Meis–Prep dimer and Meis–Prep–Pbx trimers preferentially bind the octameric or decameric (5'-TGANTGACAG-3') sequences rather than the hexameric motif¹³.

Meis1 directly regulates the transcription of the *Pax6* gene during lens morphogenesis via direct binding to its promoter region¹⁴. Like other TALE-HD TFs⁷, Meis1 directly interacts with TFs of other families, such as posterior Hox11–13 proteins¹⁵ and Pbx–Hox complexes^{16,17}, to allow combinatorial recognition of biologically meaningful binding sites. A structural study revealed that the TALE-HD of Meis1 (Meis1-HD) interacts with its consensus sequence on DNA through its $\alpha 3$ helix, in addition to electrostatic interactions with the phosphate backbone via the L1- $\alpha 1$ region (Fig. 1b)¹². However, to fully understand the molecular mechanism of the target search of Meis1-HD in genomic DNA, both structural and thermodynamic investigations into the non-specific binding state of Meis1-HD are required.

Single nucleotide mutations (SNMs) in enhancers or promoters containing the TF binding sites are able to perturb the interaction with TFs^{18–21} and can lead to various diseases^{22,23}. SNMs at the Meis1-binding site were shown to affect the binding affinity of Meis1 and are associated with diseases such as poor survival in Chinese gastric cancer patients²⁴ and hepatitis C virus-related liver disease^{23,25}. Although the SNMs disrupt the intermolecular interaction between a TF and its target DNA, a TF is usually still able to retain substantial affinity for this mutated region due to the remaining interactions. In order to explain the origin of diseases caused by SNMs at TF binding sites, the structural features of the TF complexed with various SNM-containing DNA sequences should be studied and compared with those of the TF–consensus DNA complex.

Here, we describe the structural features and the thermodynamic and kinetic properties of the non-specific DNA-binding state of the Meis1-HD, studied using NMR spectroscopy and isothermal titration calorimetry (ITC), and compare them with those of its specific binding state. The notable features of the non-specific binding state suggest a facilitated target search mechanism to explain how Meis1-HD is able to quickly and efficiently locate its specific binding site on DNA. We also investigated the structure of the Meis1-HD in complex with wild-type (WT) and various SNM-containing consensus sequences using NMR and molecular dynamics (MD) simulations. These analyses showed what kinds of structural distortions in the Meis1-HD–DNA complexes were induced by various SNMs in the consensus sequence, resulting in decreased DNA-binding affinity. This study could elucidate the detailed molecular mechanism of how Meis1-HD discriminates SNMs

from its consensus sequence. This study also leads to valuable insights into the molecular basis of the low specificity of DNA binding by Meis1.

Results

ITC data for dual binding of Meis1-HD to target DNA

The intermolecular interaction of Meis1-HD with a duplex decamer containing the consensus sequence (WT meisDNA, Fig. 1c) was thermodynamically evaluated by ITC. The ITC data fit well with a two-site binding model rather than a one-site binding model (Supplementary Fig. 1 and Fig. 1d). This analysis indicated that one Meis1-HD molecule strongly binds to the central consensus region of WT meisDNA with a $K_{d,SB}$ of 15.2 ± 2.8 nM to form the specific DNA-binding complex [(MI)^{SB}:DNA complex], and a second molecule makes non-specific contact with the remaining part of the DNA duplex with a $K_{d,NB}$ of 2.55 ± 0.64 μ M to form a 2:1 DNA-binding complex [(MI)₂:DNA complex] (Fig. 1e). The ITC data analysis produced the thermodynamic parameters for the specific and non-specific DNA-binding events (Table 1), which yielded overall free energy changes of $\Delta G^{\circ}_{SB} = -10.67 \pm 0.11$ kcal/mol and $\Delta G^{\circ}_{NB} = -7.64 \pm 1.02$ kcal/mol, respectively. These data indicated that the non-specific binding process is less favorable than the specific binding to consensus DNA, even though it is slightly entropically more favorable (Table 1).

Conformational features of WT meisDNA in the specific binding mode

The imino proton resonances of WT meisDNA (Fig. 1f) were used to study the conformational change of DNA upon binding to ¹⁵N-labeled Meis1-HD (Supplementary Fig. 2). Based on the dual DNA-binding model (Fig. 1e), the relative population of each complex as a function of the [MI]_t/[DNA]_t ratio, where [MI]_t and [DNA]_t are total concentrations of Meis1-HD and WT meisDNA, respectively, could be calculated from equations described in the Supplementary Information. When [MI]_t/[DNA]_t < 1.0, WT meisDNA exhibits a mixture of two conformational states, free DNA and the (MI)^{SB}:DNA complex (Fig. 1g). As the [MI]_t/[DNA]_t ratio is increased, the peak intensities for the T1, G2, and T5' imino protons significantly decrease until they completely disappear at [MI]_t/[DNA]_t = 0.7 (Supplementary Fig. 2), demonstrating the instability of the T1·A1', G2·C2', and A5·T5' base-pairs upon specific binding to Meis1-HD.

In order to clarify these base-pair instabilities of WT meisDNA caused by the specific binding to Meis1-HD, the hydrogen exchange rate constants (k_{ex}) of the imino protons in the WT meisDNA complexed with Meis1-HD at various [MI]_t/[DNA]_t ratios were determined at 35 °C (Supplementary Fig. 3 and Supplementary Table 1). The k_{ex} value of the T3' proton slightly increases as the [MI]_t/[DNA]_t ratio is increased up to 0.4, indicating that the A3·T3' base-pair is unusually stabilized in the specific binding complex (Fig. 1h). However, no significant change of exchange was observed for the T1, G2, G4', and T5' imino protons (Fig. 1h). Interestingly, the T3' and G4' imino resonances exhibited significant chemical shift changes as the [MI]_t/[DNA]_t ratio

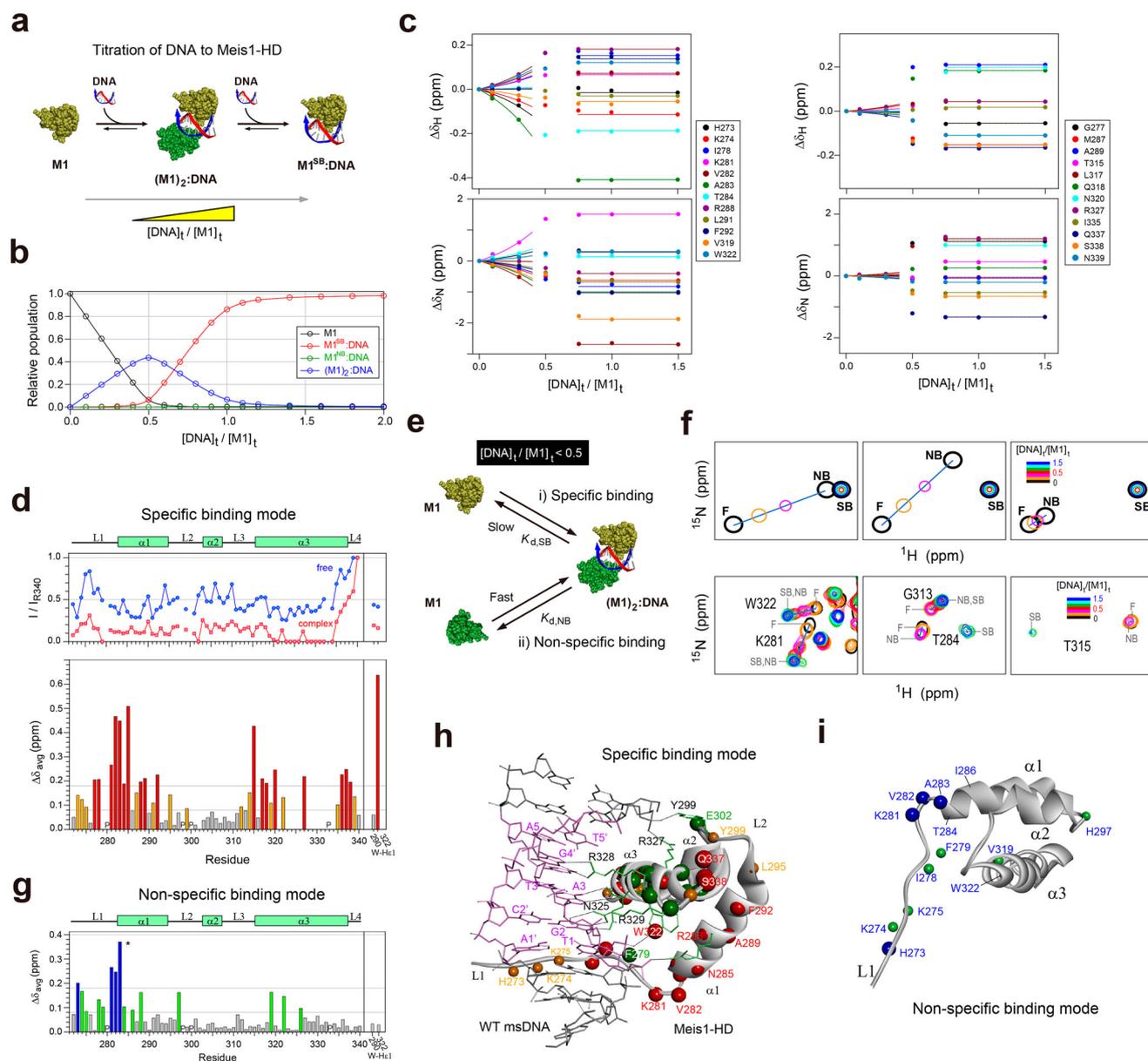


Fig. 2 | Specific and non-specific DNA-binding modes of Meis1-HD. **a** Binding scheme of Meis1-HD to DNA as the $[DNA]_t/[M1]_t$ ratio is increased. **b** Relative populations of free Meis1-HD (black) and Meis1-HD–DNA complexes ($M1^{SB}$:DNA (red), $M1^{NB}$:DNA (green), or $(M1)_2$:DNA (blue)) as a function of the $[DNA]_t/[M1]_t$ ratio. **c** (Upper) 1H and (lower) ^{15}N chemical shift change of the amide signals, which showed significant movements (left) or not (right) at $[DNA]_t/[M1]_t \leq 0.5$, of Meis1-HD upon titration with WT meisDNA as a function of the $[DNA]_t/[M1]_t$ ratio. **d** (Upper) the relative intensities (I/I_{R340}), where I_{R340} is the intensity of the R340 amide signal, of amide and W-H ϵ 1 signals of Meis1-HD in the free (blue) and specific binding complex with WT meisDNA (red). (Lower) the $\Delta\delta_{avg}$ values of Meis1-HD upon specific binding to WT meisDNA. **e** Specific and non-specific DNA-binding modes in the $(M1)_2$:DNA complex at $[DNA]_t/[M1]_t \leq 0.5$. **f** (Upper) the simulated

^{15}N -HSQC cross-peaks during titration with DNA in the cases of (left) $\delta_{Free} \neq \delta_{NB} - \delta_{SB}$, (middle) $\delta_{Free} \neq \delta_{NB} \neq \delta_{SB}$, and (right) $\delta_{Free} - \delta_{NB} \neq \delta_{SB}$. The letters, F, SB, and NB indicate the free, specific binding, and non-specific binding states, respectively. (Lower) change of $^1H/^{15}N$ -HSQC peaks of (left) K281 and W322, (middle) T284 and G313, and (right) T315 of Meis1-HD during titration with WT meisDNA. **g** The $\Delta\delta_{avg}$ values of Meis1-HD upon non-specific binding to WT meisDNA. Mapping the location of the residues with large $\Delta\delta_{avg}$ onto the crystal structure of Meis1-HD–DNA complex (PDB ID: 5BNG)¹² for **h** the specific and **i** non-specific DNA-binding modes. The colors used to illustrate the $\Delta\delta_{avg}$ are red and blue, >0.18 ppm; orange and pale green, 0.08–0.18 ppm (the same color coding is used in **d** and **g**). Green spheres in **(h)** indicate residues whose cross-peaks disappeared upon binding to DNA.

increased up to 1.0 (Fig. 1i). In addition, the T3' imino resonance also exhibited severe line-broadening (Fig. 1i).

Structural features of Meis1-HD in the specific DNA-binding mode

In the DNA-binding scheme during titration of WT meisDNA into ^{15}N -labeled Meis1-HD (Fig. 2a), most Meis1-HD molecules were involved in an $(M1)^{SB}$:DNA complex at $[DNA]_t/[M1]_t > 1.0$ (Fig. 2b) and thus no amide signals exhibited changes at $0.75 \leq [DNA]_t/[M1]_t \leq 1.5$, although

some showed significant movements at $[DNA]_t/[M1]_t \leq 0.5$ (Fig. 2c). The chemical shift perturbation (CSP) data at $[DNA]_t/[M1]_t = 1.5$ provide the 1H and ^{15}N chemical shift (δ_{SB}) for each amide signal in the specific binding mode (Supplementary Fig. 4 and Supplementary Table 2). In the specific binding complex, most amide resonances of the $\alpha 3$ helix (N321, F323–A326, R328–Q332, M334) disappeared altogether (Fig. 2d). In addition, the remaining residues of the L3- $\alpha 3$ -L4 region showed $\Delta\delta_{avg}$ values larger than 0.08 ppm (Fig. 2d). These results indicate that Meis1-HD interacts with WT meisDNA through its

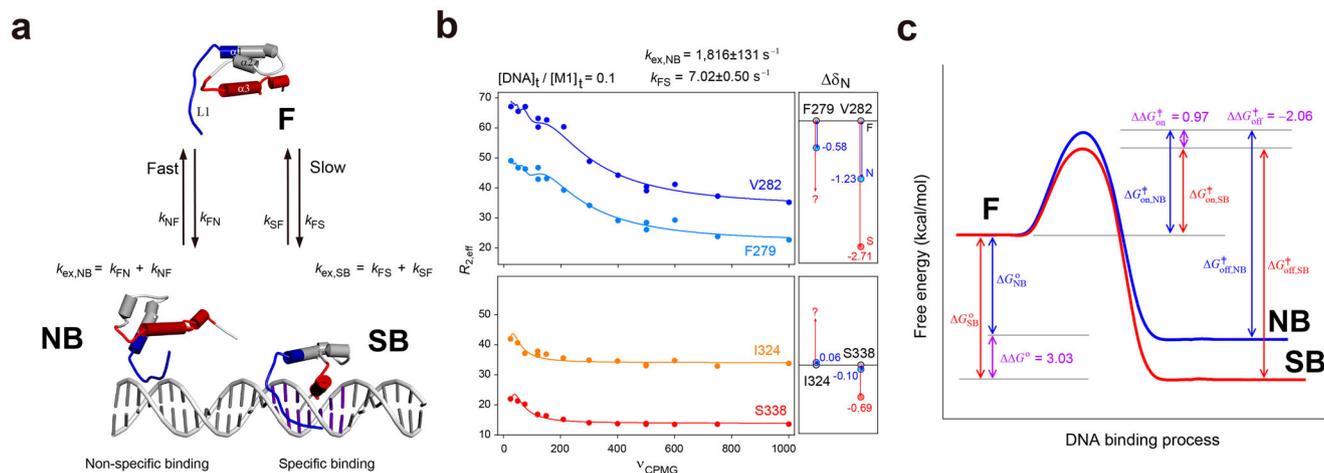


Fig. 3 | CPMG-RD study on Meis1-HD bound to DNA. **a** Representative scheme for the non-specific and specific binding modes of Meis1-HD. The letters F, NB, and SB indicate the free, non-specific binding, and specific binding states of Meis1-HD, respectively. **b** ^{15}N CPMG-RD data for the representative residues: (upper) F279 and

V282 (L1- α 1 region) and (lower) I324 and S338 (α 3-L4 region) of Meis1-HD–WT meisDNA complex determined at 800 MHz. Solid lines indicate the global best-fit for the CPMG data using Eq. 5. **c** Quantitative description of the energy landscape of the specific (red) and non-specific (blue) binding of Meis1-HD to WT meisDNA.

α 3 helix, as reported in the previous crystal structure study¹². Interestingly, most amide resonances of the L1- α 1 region (H273–F292) also had $\Delta\delta_{\text{avg}} > 0.08$ ppm (Fig. 2d). These significant chemical shift changes arise from the electric field caused by the negatively charged phosphate backbone in the minor groove of meisDNA, which participates in intermolecular H-bonding with the positively charged residues in the L1- α 1 region.

Structural features of Meis1-HD in the non-specific DNA-binding mode

To further clarify the chemical shift perturbation results, the $^1\text{H}/^{15}\text{N}$ -HSQC spectra of Meis1-HD were acquired as a function of the $[\text{DNA}]_t/[\text{M}1]_t$ ratio (Supplementary Fig. 4c). Some amide signals exhibited slow exchange between the free and DNA-bound states of Meis1-HD on the NMR time scale (Supplementary Fig. 4c and right panel of Fig. 2c). Interestingly, certain amide signals (left panel of Fig. 2c) (e.g., K281 and W322 in Fig. 2f) were strong at each titration point, indicating a fast exchange of DNA binding. Our dual DNA-binding model of Meis1-HD (Fig. 2a) can explain these unusual phenomena of the NMR titration data, which depend on both specific binding and non-specific binding in the $(\text{M}1)_2$:DNA complex. Thus, the ^1H and ^{15}N chemical shift (δ_{NB}) for each amide signal in the non-specific binding mode was indirectly determined from the titration curve at $[\text{DNA}]_t/[\text{M}1]_t < 0.5$ as described in the “Methods” section (Fig. 2c and Supplementary Table 2).

When residues have a relation that the chemical shifts for the non-specific and specific binding modes are similar to each other but differ from that of the free state ($\delta_{\text{Free}} \neq \delta_{\text{NB}} - \delta_{\text{SB}}$) like K281 and W322, their amide signals showed significant movements upon titration with WT meisDNA (left in Fig. 2f). Interestingly, when residues have different chemical shifts in all three states (that is, $\delta_{\text{Free}} \neq \delta_{\text{NB}} \neq \delta_{\text{SB}}$) like T284, their amide signals move toward different orientations compared to the position of the specific binding state (middle panel of Fig. 2f). Residues exhibiting these two kinds of patterns are listed in the left panel of Fig. 2c. In contrast, some residues, such as T315, exhibited little chemical shift changes at $[\text{DNA}]_t/[\text{M}1]_t < 0.5$, but they had significant changes at $[\text{DNA}]_t/[\text{M}1]_t \geq 1.0$ (right in Fig. 2f). Similar results were observed for residues having a relation of $\delta_{\text{Free}} - \delta_{\text{NB}} \neq \delta_{\text{SB}}$, and these are listed in the right panel of Fig. 2c.

Surprisingly, all amide signals in the L3- α 3-L4 region, except V319, W322, and A326, had $\Delta\delta_{\text{avg}}$ values between the free and non-specific

binding states smaller than 0.08 ppm, in contrast to the specific binding complex (Fig. 2g). However, most amide residues in the L1- α 1 region underwent $\Delta\delta_{\text{avg}} > 0.08$ ppm upon non-specific binding (Fig. 2g). These results indicated that the structural features of Meis1-HD in the non-specific binding mode are distinct from those of the specific binding complex. Taken together, we can summarize that: (i) when $[\text{DNA}]_t/[\text{M}1]_t > 1$, Meis1-HD forms the specific DNA-binding complex, which is similar to the crystal structure of the Meis1-HD–DNA complex (Fig. 2h); and (ii) when $[\text{DNA}]_t/[\text{M}1]_t < 0.5$, Meis1-HD forms the $(\text{M}1)_2$:DNA complex containing both the specific binding mode and the non-specific binding mode, in which the protein interacts with DNA through mainly the L1- α 1 region (Fig. 2i).

In the imino proton spectra of the WT meisDNA–Meis1-HD complex at $[\text{M}1]_t/[\text{DNA}]_t > 1.0$ (Supplementary Fig. 2), the chemical shift changes in the T3' and G4' imino resonances and the line-broadening of T3' imino resonances were not detected (Fig. 1i), meaning that the non-specific binding of Meis1-HD did not affect the conformational features of the A3-T3' and C4-G4' base-pairs.

Dynamics of specific and non-specific DNA binding of Meis1-HD

Next, in order to probe these two kinds of conformational exchange behavior of Meis1-HD in the complex with WT meisDNA, the rate constants for association and dissociation of Meis1-HD with WT meisDNA were determined using ^{15}N backbone amide Carr-Purcell-Meiboom-Gill relaxation dispersion (CPMG-RD) experiments at $[\text{DNA}]_t/[\text{M}1]_t = 0.1$. Assuming a two-state model of conformational exchange, the CPMG data set for 26 amide protons did not fit well with a single k_{ex} value using Eq. S34 ($k_{\text{ex}} = 1656 \pm 82 \text{ s}^{-1}$, $\chi^2 = 3.9858$, $R^2 = 0.9717$) or using Eq. S33 ($k_{\text{ex}} = 14.3 \pm 0.5 \text{ s}^{-1}$, $\chi^2 = 8.2549$, $R^2 = 0.9412$) (Supplementary Fig. 5). According to Fig. 2b, Meis1-HD exhibits only two conformational states, M1 (free) and $(\text{M}1)_2$:DNA, at $[\text{DNA}]_t/[\text{M}1]_t = 0.1$. We can ignore the chemical exchange between the $(\text{M}1)^{\text{SB}}$:DNA and $(\text{M}1)^{\text{NB}}$:DNA states, because these two states rarely exist under this condition (Fig. 2b). Thus, we considered the pseudo-three-state model for DNA binding of Meis1-HD, in which the fast exchange for non-specific binding and the slow exchange for specific binding are both modeled as two-state processes occurring independently of each other (Fig. 3a). The CPMG data were globally fitted using Eq. 5 (see section “Methods”) to obtain a $k_{\text{ex,NB}}$ of $1816 \pm 131 \text{ s}^{-1}$ for non-specific binding and a k_{FS} of $7.02 \pm 0.50 \text{ s}^{-1}$ for specific binding ($\chi^2 = 2.7122$, $R^2 = 0.9824$) (Supplementary Fig. 5 and Fig. 3b). The CPMG

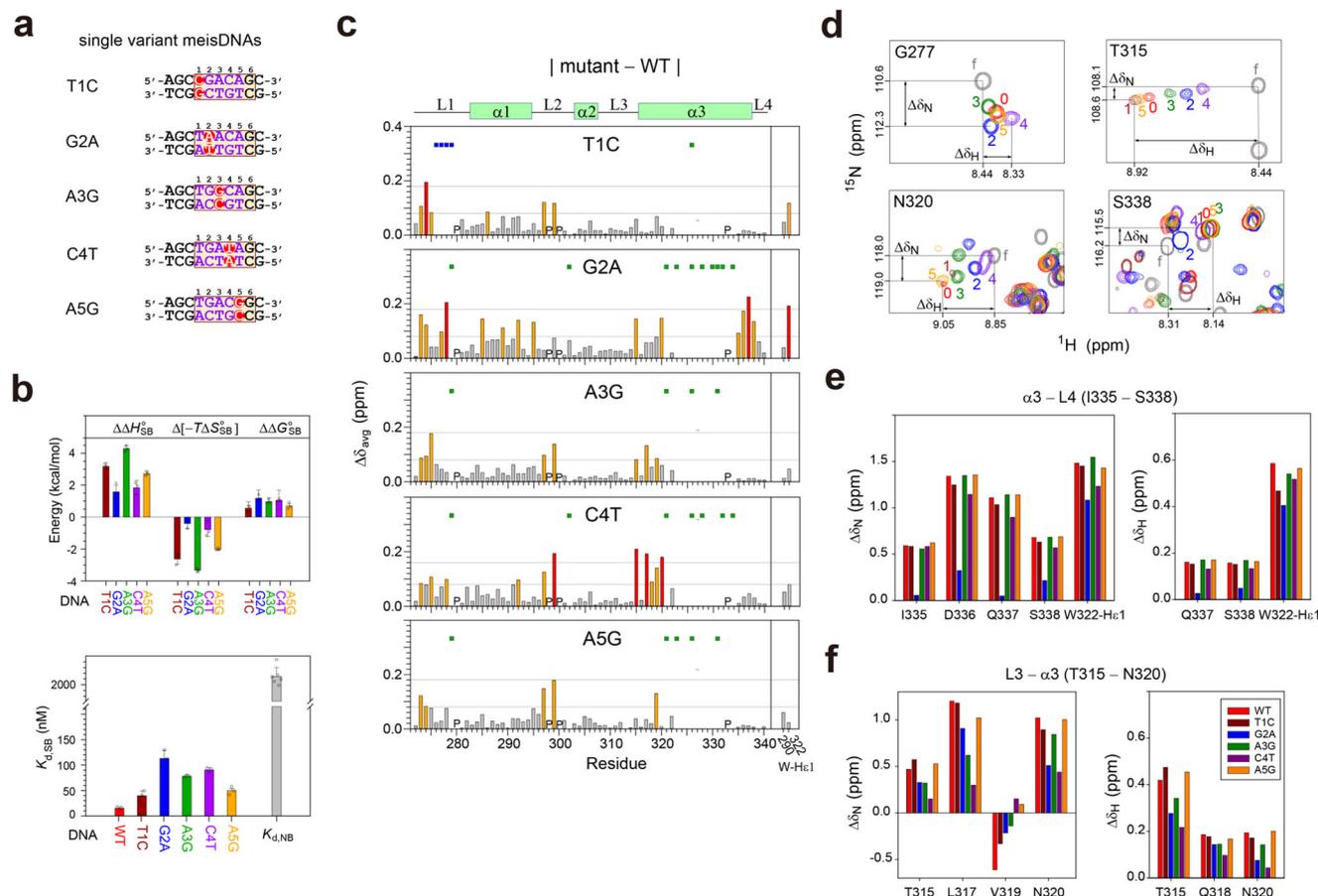


Fig. 4 | Specific binding mode of Meis1-HD to single-variant meisDNAs.

a Sequence context of single-variant meisDNAs. **b** (Upper) differences in the thermodynamic parameters between the variant and WT meisDNAs for specific binding to Meis1-HD. (Lower) dissociation constants (K_d) of Meis1-HD for specific binding to WT and variant meisDNAs and non-specific binding to meisDNA. The thermodynamic parameters were measured three times and data are presented as mean values \pm standard deviation. **c** The $\Delta\delta_{avg}$ values of Meis1-HD bound to T1C,

G2A, A3G, C4T, and A5G meisDNAs relative to the Meis1-HD-WT meisDNA complex. **d** Comparison of the $^1\text{H}/^{15}\text{N}$ -HSQC peaks of G277, T315, N320, and S338 amide protons of Meis1-HD in the free state (gray, f) and in complex with WT (red, 0), T1C (brown, 1), G2A (blue, 2), A3G (dark green, 3), T4C (purple, 4), and A5G (orange, 5) meisDNAs. ^{15}N and ^1H chemical shift changes of amide signals in the **(e)** $\alpha3$ -L4 and **(f)** L3- $\alpha3$ regions of Meis1-HD upon binding to WT and variant meisDNAs.

profile for F279 and V282, which showed significant chemical shift changes upon non-specific DNA binding ($\Delta\delta_{FN} > 0.5$ ppm), included both a fast exchange contribution at high ν_{CPMG} and slow exchange oscillation at low ν_{CPMG} (Fig. 3b). Similar results were observed for the amide signals in the L1- $\alpha1$ region (Supplementary Fig. 5). However, the amide resonances in the L3- $\alpha3$ -L4 region, such as I324 and S338, exhibited only slow exchange patterns at low ν_{CPMG} (Fig. 3b), because they were minimally involved in non-specific DNA binding ($\Delta\delta_{FN} \leq 0.1$ ppm). The CPMG-RD results were consistent with the structural features of the (MI)₂:DNA complex, determined by the CSP study.

The $k_{ex,NB}$ was used to calculate a k_{FN} of 200 ± 14 s⁻¹ and a k_{NF} of 1616 ± 117 s⁻¹ for non-specific binding. Similarly, the k_{FS} was used to calculate a $k_{ex,SB}$ of 63.82 ± 4.55 s⁻¹ and a k_{SF} of 56.80 ± 4.05 s⁻¹ for specific binding. These results indicate that Meis1-HD has a 28-fold slower apparent dissociation rate for the specific binding complex ($k_{off,SB} = k_{SF}$) than for the non-specific binding complex ($k_{off,NB} = k_{NF}$). This phenomenon can be expressed by the activation energy difference for dissociation ($\Delta\Delta G_{off}^\ddagger = -2.06$ kcal/mol) of Meis1-HD between the two binding states (Fig. 3c). Next, the apparent association rates for the specific ($k_{on,SB}^a = 3.74 \times 10^9$ M⁻¹s⁻¹) and non-specific binding ($k_{on,NB}^a = 0.62 \times 10^9$ M⁻¹s⁻¹) were calculated from the $K_{d,SB}$ and $K_{d,NB}$ values determined from the ITC data, meaning that $\Delta\Delta G_{on}^\ddagger = 0.97$ kcal/mol (Fig. 3c). From these data, we concluded

that Meis1-HD associates six times more rapidly with the consensus DNA sequence compared to non-specific contact with a random region of DNA, which can be considered as a facilitated binding of Meis1-HD to its target DNA.

ITC data for binding of Meis1-HD to single-variant meisDNAs

Meis1-HD selectively binds to the sequence, 5'-T-G-A-C-A-(G)-3'¹⁰⁻¹². To achieve this selectivity, Meis1-HD must distinguish the consensus sequence from various kinds of mutated sequences including SNMs. Here, we prepared five variant meisDNAs, which contain five SNMs in the consensus region (Fig. 4a), in order to understand the molecular basis of the sequence discrimination of Meis1-HD during DNA binding. The ITC data for these variants could be fitted with a one-site binding model, but the stoichiometric parameters (n) were much larger than one as well as significantly smaller than two (Supplementary Fig. 1). Instead, these ITC data successfully fitted with a two-site binding model to obtain two stoichiometric parameters (n) in the range of 0.98-1.03 (Supplementary Fig. 1), suggesting the two binding modes shown in Fig. 1e. Interestingly, the non-specific binding of Meis1-HD to these variants showed binding affinities as well as thermodynamic parameters similar to WT meisDNA (Table 1). These results indicated that the non-specific binding of Meis1-HD to a 10-bp DNA was not affected by a SNM of the consensus DNA. In the case of the specific binding, Meis1-HD had $K_{d,SB}$ values in the range of 40-113 nM for the

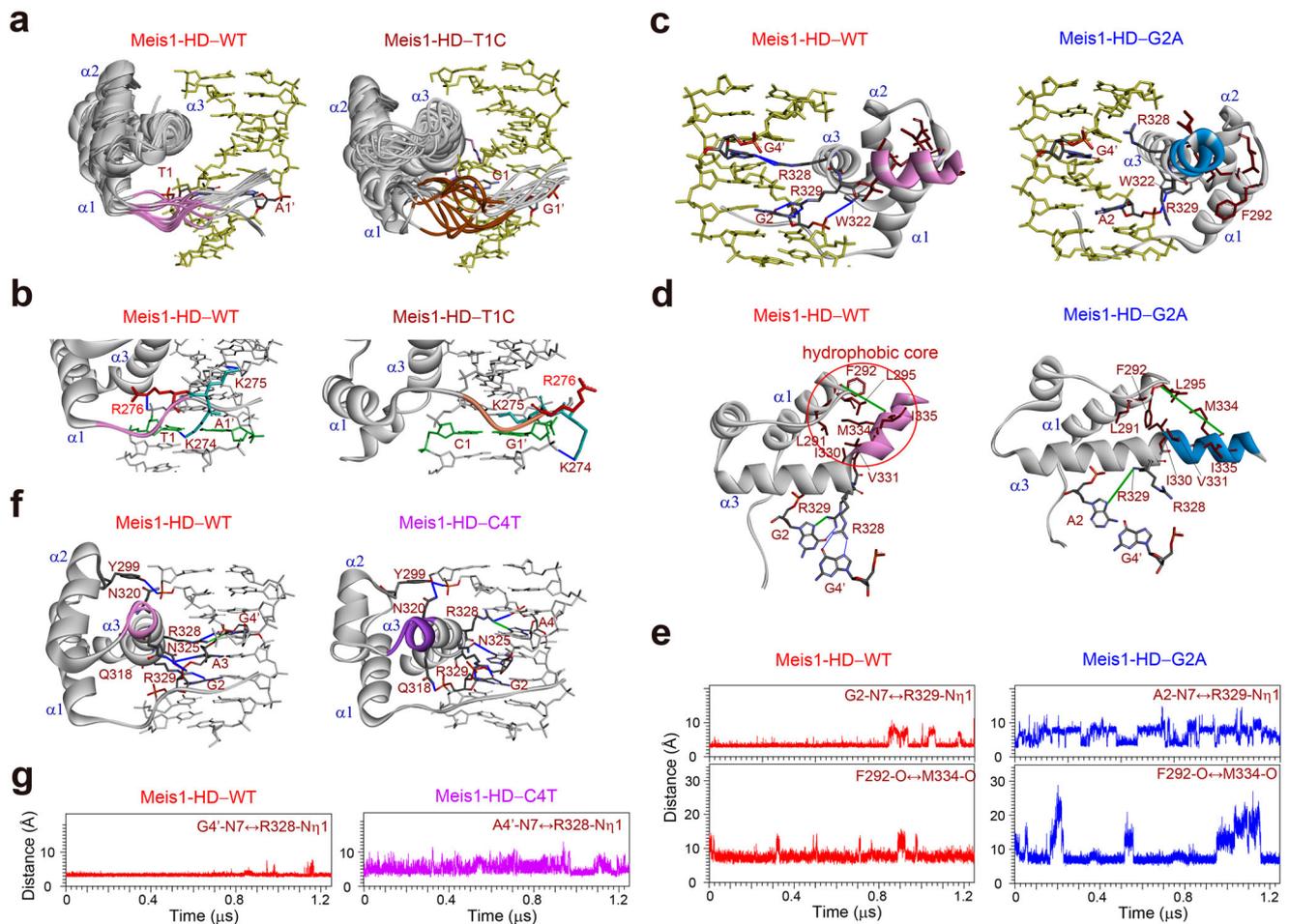


Fig. 5 | MD simulations of Meis1-HD in complex with WT or variant meisdNAs. **a** Superimposition of Meis1-HD in complex with (left) WT and (right) T1C meisdNAs. **b** Intermolecular H-bonding interaction of the L1 loop region in the (left) Meis1-HD-WT and (right) Meis1-HD-T1C complexes. **c** Intermolecular H-bonding interaction of residues W322, R328, and R329 in the (left) Meis1-HD-WT and (right) Meis1-HD-G2A complexes. **d** Hydrophobic core created between the $\alpha 1$ and $\alpha 3$ helices in the (left) Meis1-HD-WT and (right) Meis1-HD-G2A complexes. **e** Distances of (upper) X2-N7 \leftrightarrow R329-N η 1 and (lower) F292-O \leftrightarrow M334-O, where X2 is G2 or A2, in the (left) Meis1-HD-WT and (right) Meis1-HD-G2A complexes as a

function of the simulation time. The distances are indicated as green solid lines in (d). **f** Intermolecular H-bonding interaction of the $\alpha 3$ helix in the (left) Meis1-HD-WT and (right) Meis1-HD-C4T complexes. **g** Distances of X4'-N7 \leftrightarrow R328-N η 1, where X4' is G4' or A4', in the (left) Meis1-HD-WT and (right) Meis1-HD-C4T complexes as a function of the simulation time. These distances are indicated as green solid lines in (f). The representative structures are the centroids of the largest clusters among 10 conformational clusters obtained with the k-means clustering method. The largest clusters of WT, TIC, G2A, and C4T meisdNAs occupy 26.6%, 37.4%, 38.1%, and 25.9% of the trajectory, respectively.

variants (Fig. 4b and Table 1), indicating that the specific binding to single variants was still at least 20 times stronger than the non-specific binding.

Specific binding of Meis1-HD to the T1C variant

The T1C variant had significantly reduced enthalpy for the specific binding, with a $\Delta\Delta H^{\circ}_{SB} (= \Delta H^{\circ}_{SB,T1C} - \Delta H^{\circ}_{SB,WT})$ of 3.18 kcal/mol (Fig. 4b). Instead, this specific binding increased entropy with a $\Delta[-7\Delta S^{\circ}_{SB}] (= [-7\Delta S^{\circ}_{SB,T1C}] - [-7\Delta S^{\circ}_{SB,WT}])$ of -2.62 kcal/mol (Fig. 4b). Thus, Meis1-HD has only a 2.6-fold larger $K_{d,SB}$ value for the T1C than WT meisdNA (Fig. 4b). These data indicated that the T1C SNM only slightly affected the DNA-binding affinity of Meis1-HD, because this DNA-binding process is entropically favorable, even though it is enthalpically unfavorable.

The T1C variant caused significant CSPs on residues in the L3- $\alpha 3$ -L4 as well as the L1- $\alpha 1$ region at $[DNA]_t/[M]_t = 1.5$, like WT meisdNA (Supplementary Figs. 6a and 7a). Interestingly, the only significant $\Delta\delta_{avg}$ differences between the complexes with WT and T1C meisdNAs were observed in the residues in the L1 and TALE regions (Fig. 4c). The most striking feature was that the amide resonances of residues

R276-F279 disappeared upon binding to T1C, in contrast to WT meisdNA (e.g., G277 in Fig. 4d), indicating unusual flexibility of the L1 loop.

The MD simulation study of the Meis1-HD-T1C complex proved this high structural flexibility of the L1 loop region (Fig. 5a). From the cluster analysis of the MD trajectories, four distinct conformational clusters of the L1 loop were observed, whereas the conformational clusters of the WT complexes were well superimposed, indicating low flexibility (Fig. 5a, pink and brown colors indicate residues R276-F279). MD simulation revealed that the change of T1-A1' to a C1-G1' base-pair disrupted the H-bonding interaction of L1 with DNA (K274-H ζ \leftrightarrow T1-O2 or R276-H η \leftrightarrow T1-O2) and then caused the flexibility of L1 (Fig. 5b). The H-bond of K274-H ζ \leftrightarrow T1-O2 or R276-H η \leftrightarrow T1-O2 is observed 40% of the MD trajectories of WT meisdNA in average, while the corresponding H-bonds were only observed in <6% of a trajectory among three independent trajectories (Supplementary Tables 3 and 4). Additionally, the root mean square fluctuation (RMSF) analysis of Meis1-HD residues demonstrates that binding to T1C increases the flexibility of the L1 loop significantly (Supplementary Fig. 8a and Supplementary Table 5). These structural features of the Meis1-HD-T1C complex are

consistent with the thermodynamic properties of the specific binding to TIC (enthalpically unfavorable but entropically favorable) (Fig. 4b).

Specific binding of Meis1-HD to the G2A variant

The G2A variant showed a $\Delta\Delta H_{\text{SB}}^{\circ}$ value of 1.58 kcal/mol for the specific binding (Fig. 4b). Unlike TIC, the G2A did not show entropically favorable behavior ($\Delta[-T\Delta S_{\text{SB}}^{\circ}] = -0.89$ kcal/mol) (Fig. 4b). Thus, the specific binding affinity of Meis1-HD for G2A is 7.4-fold weaker than for WT *meisDNA* (Fig. 4b). In the Meis1-HD–WT *meisDNA* complex, the side chains of R329 and R328 form H-bonding interactions with the G2 and G4' bases of WT *meisDNA*, respectively (Fig. 5c). However, these intermolecular H-bonding interactions are completely disrupted in the Meis1-HD–G2A complex (Fig. 5c), confirmed by comparison of the A2-N7 \leftrightarrow R329-H η 1 (3–9 Å) vs. G2-N7 \leftrightarrow R329-H η 1 (<4 Å) distances in the MD simulation (Fig. 5e and Supplementary Fig. 9). The H-bonds between the G2-base and the side chain of R329 were maintained in more than 60% of the trajectory in the Meis1-HD–WT *meisDNA* complex, while the corresponding H-bond was observed <1.0% of the trajectory in the Meis1-HD–G2A complex (Supplementary Fig. 9). This is consistent with the distinct CSP results caused by G2A compared to WT and other variant *meisDNAs* (Fig. 4c). Most amide signals in the α 3 helix were observed in the Meis1-HD–G2A complex (Supplementary Figs. 6b and 7a), whereas they disappeared in the Meis1-HD–WT *meisDNA* complex (Fig. 2d).

Surprisingly, residues I335–S338 in the α 3-L4 region had $\Delta\delta_{\text{avg}}$ differences >0.08 ppm between the complexes with WT and G2A *meisDNA* (Fig. 4c), although this region is far from the mutated A2-T2' base-pair (Supplementary Fig. 7c). For example, residue S338 in L4 exhibited a smaller upfield shift of its amide resonance upon binding to G2A than WT and other variants (Fig. 4d). Similarly, the amide resonances of residues I335–Q337 showed less movement upon binding to G2A compared to other complexes (Fig. 4e). In the Meis1-HD–WT *meisDNA* complex, this α 3-L4 region exhibited an L-shape kinked structure to form a hydrophobic core with the α 1 helix (Fig. 5d), confirmed by the F292-O \leftrightarrow M334-O distance (<9 Å) in the MD simulation (Fig. 5e). However, in the MD simulation for the Meis1-HD–G2A complex, this distance frequently became longer than 13 Å (Fig. 5e), indicating the straight structure of the α 3 helix (Fig. 5d). The RMSF analysis also confirms that binding to G2A reduces the overall movement of the α 3 helix (Supplementary Fig. 8b).

Specific binding of Meis1-HD to the C4T variant

The C4T variant showed $\Delta\Delta H_{\text{SB}}^{\circ}$ and $\Delta[-T\Delta S_{\text{SB}}^{\circ}]$ values of 1.83 and –0.78 kcal/mol, respectively, for the specific binding (Fig. 4b). Thus, the $K_{\text{d,SB}}$ value of Meis1-HD for C4T is 5.9-fold larger than that for WT *meisDNA* (Fig. 4b). This is consistent with the fact that the intermolecular H-bonding interaction of the R328 side chain is disrupted by the C4T change (Fig. 5f).

Unlike G2A, this complex showed a similar CSP pattern to WT *meisDNA* in the residues I335–S338 (that is, $\Delta\delta_{\text{avg}} < 0.08$ ppm) (Fig. 4c). Instead, the residues in the L3- α 3 region (residues I315–N320) had $\Delta\delta_{\text{avg}}$ differences >0.08 ppm (Fig. 4c). For example, the C4T variant caused the least movement of the T315 and N320 amide signals among *meisDNAs* (Fig. 4d). Similar results were observed for the L317, Q318, and V319 amide resonances (Fig. 4f). In the Meis1-HD–C4T complex, the R328 side chain forms an H-bonding interaction with the A4'pT5' phosphate instead of the A4' base (Fig. 5f), confirmed by the A4'-N7 \leftrightarrow R328-H η 1 distance (3–11 Å) (Fig. 5g and Supplementary Fig. 10). The H-bond between the G4' base and side chain of R328 was observed for 29% of the trajectory in the Meis1-HD–WT *meisDNA* complex, while A4' formed H-bonds only with DNA backbone atoms in the Meis1-HD–C4T complex (Supplementary Fig. 10). In addition, the Q318 side chain forms an H-bond with the T1pG2 phosphate, whereas this H-bonded with the G2pA3 phosphate in the complex with WT *meisDNA* (Fig. 5f). As a result, in the Meis1-HD–C4T complex, the α 3

helix exhibited a V-shape structure which is kinked at a different position from the Meis1-HD–WT *meisDNA* complex (Fig. 5f).

Specific binding of Meis1-HD to the A3G and A5G variants

In the ITC data, the A3G and A5G variants showed similar thermodynamic properties to TIC (Fig. 4b and Table 1). Consequently, Meis1-HD had 5.1- and 3.3-fold larger $K_{\text{d,SB}}$ values for A3G and A5G, respectively, compared to WT *meisDNA* (Fig. 4b). These two variants caused $\Delta\delta_{\text{avg}}$ differences in the residues in the L3- α 3 region with much smaller CSPs than C4T (Fig. 4c). RMSF analysis revealed that the Meis1-HD showed extreme overall flexibility upon binding to A3G *meisDNA* (Supplementary Table 7 and Supplementary Fig. 8c). In the Meis1-HD–A5G complex, the R328-G4' H-bonding interaction exhibited highly dynamic features, confirmed by the G4'-N7 \leftrightarrow R328-H η 1 distance (Supplementary Table 10 and Supplementary Fig. 10). These MD simulation results are consistent with the thermodynamic properties of the specific binding to A3G and A5G variants (positive $\Delta\Delta H_{\text{SB}}^{\circ}$ and negative $\Delta[-T\Delta S_{\text{SB}}^{\circ}]$) (Fig. 4b).

We further investigated the sum of the fractions of H-bonds of Meis1-HD \leftrightarrow DNA backbone and Meis1-HD \leftrightarrow DNA bases observed in the MD trajectories (Supplementary Fig. 11). For all *meisDNA* complexes, except A5G, the fractions of specific interactions, H-bonds of Meis1 \leftrightarrow DNA base, are significantly reduced compared to WT *meisDNA*, while those of non-specific interactions (Meis1-HD \leftrightarrow DNA backbone) are increased. These results suggest that Meis1-HD–WT *meisDNA* interactions are optimized to the target sequence, and the target sequence mutation converts sequence-specific H-bonds to non-specific H-bonds.

Discussion

Like other HD and TALE-HD proteins, the TALE-HD of Meis1 has two DNA contact regions, the L1- α 1 region and the α 3 helix. The specific binding mode of Meis1-HD includes these two interactions, both of which are optimized to be efficiently bound to a target sequence. The previous crystal structure study revealed that Meis1-HD binds to consensus DNA (WT *meisDNA*) through these two contact regions¹². However, in this structure, most of the L1 region was missing, and thus only R288 of the α 1 helix shows the H-bonding interaction with DNA¹². Residues R329, N325, and R328 of the α 3 helix also form intermolecular H-bonding interactions with the G2, A3, and G4' bases of DNA, respectively, to recognize the consensus sequence¹².

In this complementary study, we demonstrated the distinctive structural and thermodynamic features of the non-specific DNA-binding state of Meis1-HD as an intermediate structure during target DNA recognition by Meis1-HD (Fig. 6). First, the second binding of Meis1-HD to various 10-bp DNA duplexes showed the sequence independence of their thermodynamic parameters, implying the non-specific binding mode of Meis1-HD. The binding affinities and thermodynamic parameters of five SNMs for the second DNA-binding event are similar to those of WT *meisDNA* (Table 1), strongly supporting the DNA sequence independence. The slight differences in these parameters can be ignored, and have two possible explanations: (i) the initial strong binding of Meis1-HD can affect the subsequent weak non-specific binding; (ii) the inaccuracy in the fitting of the thermodynamic parameters may arise from similarities in the two binding modes.

Second, the non-specific binding mode of Meis1-HD includes mainly the intermolecular interaction by the L1- α 1 region. The L1- α 1 region (K271–R288) contains seven positively charged residues (4 Lys and 3 Arg) (Fig. 1a). Our MD simulations revealed that, in the specific binding complex, residues K274, K275, and R276 of the L1 loop participated in H-bonding interactions with the phosphate backbone of DNA through the minor groove (Fig. 5b). Similarly, in the non-specific binding complex, the positively charged residues are able to electrostatically interact with the phosphate backbone of DNA, which is

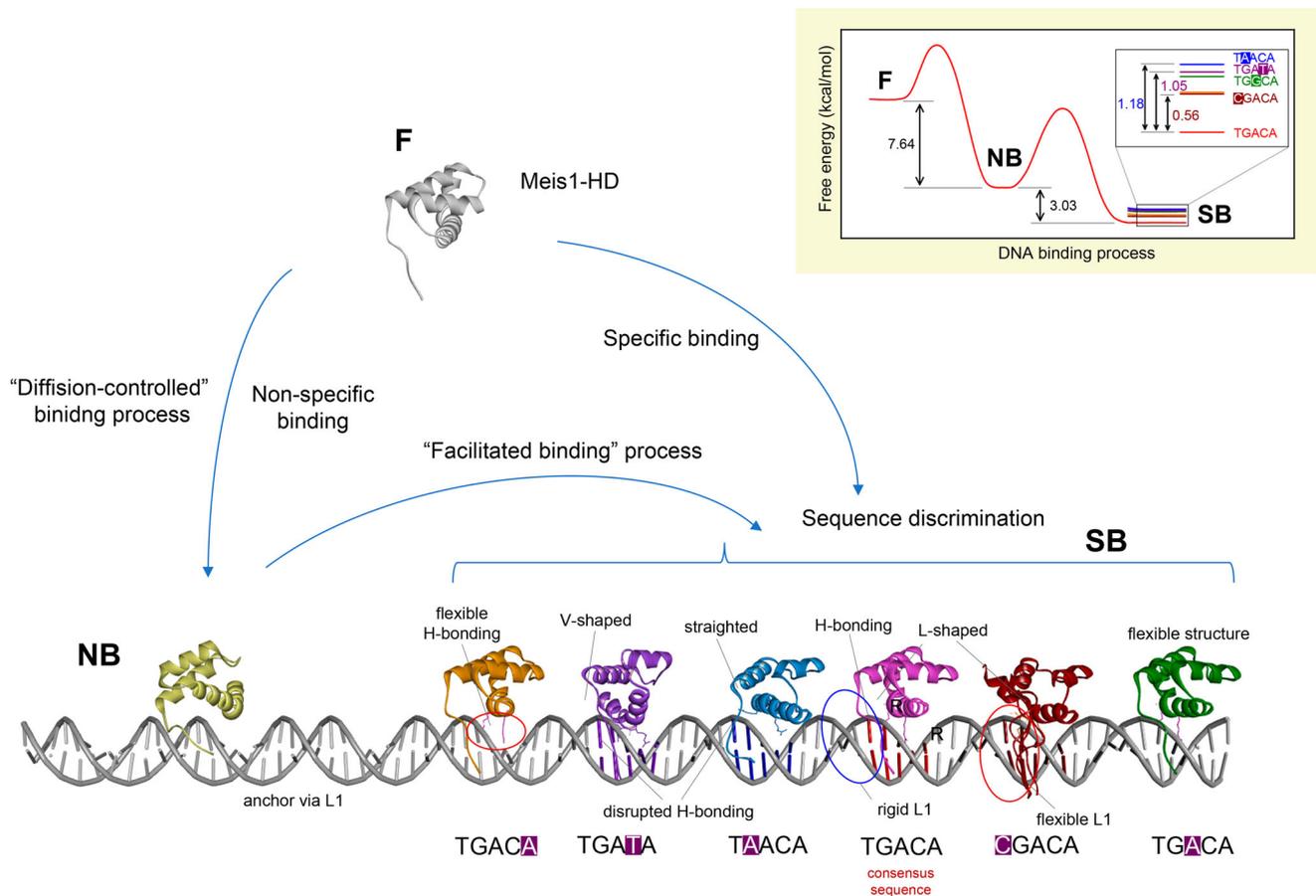


Fig. 6 | Proposed mechanism for facilitated target DNA recognition by Meis1-HD. (1) Free Meis1-HD (F) first non-specifically contacts DNA to form an anchor structure (NB) via the L1 loop, in a diffusion-controlled binding process. (2) Then, Meis1-HD rapidly translocates to the consensus DNA site by additional binding of $\alpha 3$ to form the specific binding state (SB), which is a facilitated binding process. (3)

In the specific binding state, Meis1-HD exactly recognizes the consensus sequence by the conformational distortion and flexibility features of the DNA. A quantitative description of the energy landscape of the non-specific and specific binding states of Meis1-HD to WT and single-variant meisdNA is shown in the inset.

confirmed by the significant CSP of the L1- $\alpha 1$ region (H273–F292) (Fig. 2g).

Third, the non-specific binding of Meis1-HD is the energetically preferred process to achieve the specific binding to the consensus sequence (ΔG_{NB}^0 accounts for 72% of ΔG_{SB}^0). A triple-variant, G2A/A3G/C4T, in which the H-bonding interactions with the $\alpha 3$ helix are disrupted, has an 80-fold larger $K_{d,SB}$ and significantly reduced enthalpy ($\Delta H_{SB}^0 = -5.64$ kcal/mol) compared to WT meisdNA (Supplementary Fig. 12). This $K_{d,SB}$ value became similar to the $K_{d,NB}$ of Meis1-HD (Supplementary Fig. 12). Thus, we suggest that the positively charged residues of the L1- $\alpha 1$ region might be an anchor to form the stable non-specific binding complex (Fig. 6). This intermediate complex can be easily converted to the specific binding state by H-bonding interactions of the $\alpha 3$ helix with DNA bases through the major groove to recognize the consensus sequence (Fig. 6).

Fourth, Meis1-HD associates six times more rapidly with the consensus sequence than with a random DNA (Fig. 3). In order to efficiently achieve specific target recognition, a facilitated searching process is required rather than a diffusion-controlled process, which would take longer to locate the consensus binding site within an overwhelming amount of genomic DNA. Three facilitated binding mechanisms have been previously described: 1D diffusion along the DNA, intramolecular hopping, and intersegment transfer^{1,26–30}. If we assume the non-specific contact of Meis1-HD with DNA ($k_{on,NB}^0 = 0.62 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$) is a diffusion-controlled process ($k_{diffusion} = 10^8\text{--}10^9 \text{ M}^{-1} \text{ s}^{-1}$)³¹, the fast association of the specific binding

complex ($k_{on,NB}^0 = 3.74 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$) might explain the existence of a facilitated binding process. Thus, we can suggest a two-step model for target DNA recognition of Meis1-HD: (i) Meis1-HD first non-specifically contacts DNA to form an anchored structure via the L1 loop; (ii) then, Meis1-HD is able to rapidly translocate to the consensus region via the specific binding mode, which is considered a facilitated binding process (Fig. 6).

Meis1-HD exactly recognizes the consensus sequence, 5'-T1-G2-A3-C4-A5-G6-3', through the $\alpha 3$ helix (Fig. 1b)^{10–12}. SNMs at this consensus site disrupt the interaction with Meis1, leading to related diseases^{22–25}. Transition mutations, which are interchanges of pyrimidines (T \leftrightarrow C) or of purines (A \leftrightarrow G), account for 70% of all SNMs²¹. A single C \rightarrow T transition mutation at the Meis1-binding site, which corresponds to the G2A variant in this study, was associated with poor survival in Chinese gastric cancer patients²⁴. Conversely, the A \rightarrow G mutation in the *CCL2* gene increases the Meis1-binding affinity and results in hepatitis C virus-related liver disease^{23,25}.

The studies of Meis1-HD complexed with various SNM-containing DNAs found several key structural features that play an important role in the sequence discrimination step of Meis1-HD. In contrast to the non-specific binding, the thermodynamic parameters of the strong initial binding of Meis1-HD to various 10-bp DNA duplexes clearly exhibited sequence dependence, as did the structural features of the complexes with Meis1-HD, implying the specific binding mode of Meis1-HD. Five SNMs studied here can be divided into two categories according to the thermodynamics data: (i) G2A and C4T; (ii) TIC, A3G,

and A5G (Fig. 4b). The G2A and C4T variants showed positive $\Delta\Delta H_{SB}^{\circ}$ (G2A = 1.58, C4T = 1.83 kcal/mol) and slightly negative $\Delta[-T\Delta S_{SB}^{\circ}]$ (G2A = -0.34, C4T = -0.78 kcal/mol) (Fig. 4b). In the Meis1-HD-WT meisDNA complex, R329 and R328 specifically recognize the G2 and G4' bases, respectively, via H-bonding interactions (Fig. 5c). MD simulations found that the α 3-L4 region should form an L-shape kinked structure (i.e., the active conformation) in order to maintain these interactions (Fig. 5d). This study revealed that the G2A variant is unable to interact with both the R329 and R328 side chains (Fig. 5c) and thus the α 3 helix exhibited the straight structure (the "inactive conformation") (Fig. 5d). In C4T, these interactions were also completely disrupted (Fig. 5f). Instead, R328 and Q318 side chains form H-bonding interactions with the phosphate backbone, leading to formation of a V-shape structure of the α 3 helix (another inactive conformation) (Fig. 5f). Thus, we concluded that Meis1-HD is able to distinguish the G2A (G2-C2'→A2-T2') and C4T (C4-G4'→T4-A4') SNMs through the inactive conformation of α 3 helix caused by the disruption of the H-bonding interactions of R329 and R328 (Fig. 6).

Surprisingly, the TIC, A3G, and A5G variants showed highly positive $\Delta\Delta H_{SB}^{\circ}$ values (2.72–4.29 kcal/mol) but significantly negative $\Delta[-T\Delta S_{SB}^{\circ}]$ values (-2.02 ~ -3.32 kcal/mol) (Fig. 4b). The MD simulation study revealed that, in the Meis1-HD-WT meisDNA complex, the L1 loop exhibited a rigid structure because of several H-bonding interactions with DNA (Fig. 5b). However, the TIC variant disrupts these H-bonding interactions and then the L1 loop displays highly flexible structure (Fig. 5a and Supplementary Fig. 8a). In the Meis1-HD-A3G complex, the overall region of Meis1-HD exhibited highly dynamic features (Supplementary Fig. 8c). In the case of the A5G complex, the R328 side chain forms H-bonding interactions with the G4' base with highly dynamic features (Supplementary Fig. 10). These MD simulation results are consistent with the thermodynamic properties of the specific binding to the TIC, A3G, and A5G variants (enthalpically unfavorable but entropically favorable) (Fig. 4b). Thus, we also concluded that Meis1-HD can discriminate the TIC (T1-A1'→C1-G1') and A3G (A3-T3'→G3-C3') SNMs by the highly flexible conformation of the L1 loop and overall protein, respectively, and the A5G (A5-T5'→G5-C5') SNM through the dynamic features of the H-bonding interaction of R329 (Fig. 6).

Human TALE-HD TFs have distinct target sequences for DNA binding^{10–12,32,33}. The target DNA sequence of the Prep family is the same as the hexameric sequence of the Meis family³², whereas the Pbx family preferentially binds the octameric 5'-TGATTGAT-3' sequence³³. Interestingly, their DNA-binding affinities and specificities were changed by the formation of the heterodimers or heterotrimers with other HD TFs, like Hox homeobox proteins¹³. The target DNA sequences of the Meis1 monomer is hexameric (corresponding to WT meisDNA) and octameric sequence (corresponding to C4T meisDNA)¹³. The Meis1-Prep1 and Meis1-Pbx1 dimers and the Meis1-Prep-1Pbx1 trimer bind preferentially to the octameric motif³³. Our study revealed that Meis1-HD is still able to strongly bind the SNMs with $K_{d,SB}$ values in the range of 40–113 nM (Table 1), which is only slightly weaker as measured for the canonical Meis binding site (Fig. 4b). The strong DNA-binding affinity and broad specificity of Meis1 ($K_{d,SB} \leq 113$ nM for SNMs) suggest a biological function of Meis1 as a TF-interacting protein that forms dimers or trimers with other TFs to endow the resulting assembly with an overall superior specificity for a biologically functional binding site.

In summary, we have performed thermodynamic and structural studies of TALE-HD, Meis1-HD, complexed with DNA duplexes containing the WT consensus sequence and five different SNMs. In addition to the specific binding state, in which Meis1-HD recognizes the target DNA sequences via the L1- α 1 region and the α 3 helix, we demonstrated that the non-specific binding state is achieved by the intermolecular interaction of the L1- α 1 region with the DNA. NMR dynamics studies suggested that the non-specific binding state might represent an intermediate structure in the two-step facilitated binding

model: (i) Meis1-HD first non-specifically contacts DNA to form an anchored structure; (ii) then, Meis1-HD is able to rapidly translocate to the consensus region via the specific binding state. Our structural studies also showed the key structural features that explain how Meis1-HD is able to distinguish the consensus sequence, 5'-T1-G2-A3-C4-A5-3', from other SNMs. This study concluded that Meis1-HD distinguishes the G2A and C4T mutations through the conformational features (active or inactive) of the α 3 helix and the TIC, A3G, and A5G mutations by the dynamic features (rigid or flexible) of the L1 loop and/or the α 3 helix.

Methods

Sample preparation

The DNA oligomers were purchased from M-Biotech Inc. (Seoul, Korea) and desalted using a Sephadex G-25 gel filtration column. The following DNA oligomers were utilized in this study: WT meisDNA (5'-AGCTGACAGC-3'), TIC variant (5'-AGCCGACAGC-3'), G2A variant (5'-AGCTAACAGC-3'), A3G variant (5'-AGCTGGCAGC-3'), C4T variant (5'-AGCTGATAGC-3'), A5G variant (5'-AGCTGACGGC-3') and triple-variant (5'-AGCTAGTAGC-3'). The DNA coding sequences for Meis1-HD (aa 271–340) were purchased from BIONEER Inc. (Daejeon, Korea) and cloned into *E. coli* expression vector pGEX-2T, which has an N-terminal GST tag. To produce uniformly ¹⁵N-labeled or ¹³C/¹⁵N-labeled Meis1-HD, BL21(DE3) cells were grown in an M9 minimal medium that contained 1 g/L ¹⁵NH₄Cl and/or 2 g/L ¹³C-glucose as the sole nitrogen and carbon sources. The expressed proteins were purified by a GST affinity column and a Sephacryl S-100 gel filtration column on a GE AKTA FPLC system (GE Healthcare, Chicago, IL, USA). The DNA and protein samples were dissolved in a 90% H₂O/10% D₂O NMR buffer containing 10 mM sodium phosphate (pH = 6.0) and 100 mM NaCl. The concentrations of all protein and DNA samples were measured using an N60 NanoPhotometer (Implen Inc., Germany).

ITC experiments

A Nano-ITC calorimeter (TA Instruments, DE, USA) was used for the ITC experiments (GNU). Protein and DNA samples were extensively dialyzed against 2 L of 10 mM sodium phosphate buffer (pH = 6.0) containing 100 mM NaCl. In the titration experiments, 10 μ L of 600 μ M DNA solution were added into the 75 μ M solution of Meis1-HD each time, and 25 injections were performed. Other experimental conditions were as follows: interval, 200 s; stirring speed, 350 rpm; and cell temperature, 298 K.

NMR experiments

All of the ¹H and ¹⁵N NMR experiments were performed on an Agilent DD2 700-MHz NMR spectrometer (GNU, Jinju) equipped with a triple resonance probe or Bruker Avance-Neo 900 MHz (KBSI, Ochang) and Avance-III HD 800 MHz NMR spectrometers (KIST, Seoul) equipped with a cold probe. The imino proton and ¹H/¹⁵N-HSQC spectra were obtained for complexes prepared by the addition of DNA to 0.5 mM ¹⁵N-labeled Meis1-HD at 298 K. One-dimensional (1D) NMR data were processed with Mnova 12.0.0 (Mestrelab, Spain) software, while the 2D data were processed with NMRPipe³⁴ and analyzed with NMRFAM-Sparky³⁵. External 2,2-dimethyl-2-silapentane-5-sulfonate was used for the ¹H, ¹³C, and ¹⁵N references.

¹H, ¹³C, and ¹⁵N resonance assignments for Meis1-HD were obtained from the following 3D experiments in 10% D₂O/90% H₂O containing 10 mM sodium phosphate (pH 6.0) and 100 mM NaCl: CACB(CO)NH, HNCA, HNCACB, HNCO, HN(CA)CO, HCCH-TOCSY, NOESY-¹H/¹³C-HSQC, NOESY-¹H/¹⁵N-HSQC and TOCSY-¹H/¹⁵N-HSQC. The average chemical shift differences ($\Delta\delta_{avg}$) of the amide proton and nitrogen resonances between the two states were calculated by Eq. 1:

$$\Delta\delta_{avg} = \sqrt{(\Delta\delta_H)^2 + (\Delta\delta_N/5.88)^2} \quad (1)$$

where $\Delta\delta_{\text{H}}$ and $\Delta\delta_{\text{N}}$ are the chemical shift differences of the amide proton and nitrogen resonances, respectively. The threshold values of the $\Delta\delta_{\text{avg}}$ data were determined from their standard deviation (σ) excluding any residues for which the $\Delta\delta_{\text{avg}}$ is $>3 \times \sigma$ ³⁶.

For the NMR titration experiments, both 1D imino spectra and ¹H-¹⁵N-HSQC spectra were acquired under the same buffer conditions as the 3D experiments at 298 K. The titration series involved titrating 0.5 mM of Meis1-HD with meisDNA using the following $[\text{DNA}]_t/[\text{MI}]_t$ ratios: 0, 0.1, 0.3, 0.5, 0.75, 1.0, and 1.5. Additionally, 0.5 mM WT meisDNA was titrated with Meis1-HD using the following $[\text{MI}]_t/[\text{DNA}]_t$ ratios: 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.3, and 1.5.

Hydrogen exchange rate measurements

The hydrogen exchange rate constants (k_{ex}) of the imino protons of free DNA and the DNA–Meis1-HD complexes at various $[\text{MI}]_t/[\text{DNA}]_t$ ratios were measured by the phase-modulated CLEAN chemical exchange (CLEANEX-PM)³⁷ method (see Supplementary Note 2). The CLEANEX-PM experiments employed 10 values of the spin-lock mixing time (t) ranging from 2 to 20 ms. The k_{ex} values were calculated for each cross-peak signal at each value by:

$$\frac{I(t)}{I_0} = \frac{k_{\text{ex}}}{R_{2w} - R_{2a}} (e^{-R_{2a}t} - e^{-R_{2w}t}) \quad (2)$$

where I_0 and $I(t)$ are the peak intensities of the imino proton in the CLEANEX-PM experiments at mixing times 0 and t , respectively, and R_{2a} and R_{2w} are the apparent transverse relaxation rate constants for the imino proton and water, respectively. The k_{ex} values were confirmed by the water magnetization transfer experiments^{38,39} (Supplementary Table 1).

Deconvolution of the 1H and 15N amide chemical shifts of Meis1-HD in the specific and non-specific binding states

When $[\text{DNA}]_t/[\text{MI}]_t > 1.0$, Meis1-HD exhibits only the (MI)₁:DNA complex (Fig. 2b) and thus the ¹H and ¹⁵N chemical shifts (δ_{SB}) for each amide signal in the specific binding mode were directly obtained from the ¹H/¹⁵N-HSQC spectrum at $[\text{DNA}]_t/[\text{MI}]_t = 1.5$.

When $[\text{DNA}]_t/[\text{MI}]_t < 0.5$, Meis1-HD exhibited a mixture of two conformational states, free Meis1-HD and (MI)₂:DNA complex (Fig. 2b). Because only the fast exchange process for non-specific binding affects the movement of amide signals, the ¹H and ¹⁵N chemical shifts (δ_{NB}) for each amide signal in the non-specific binding mode were determined by individually fitting the observed chemical shift changes ($\Delta\delta_{\text{obs}} = \delta_{\text{obs}} - \delta_{\text{Free}}$) at $[\text{DNA}]_t/[\text{MI}]_t < 0.5$ to Eq. 3:

$$\Delta\delta_{\text{obs}} = \frac{[(\text{MI})_2 : \text{DNA}]}{[\text{MI}] + [(\text{MI})_2 : \text{DNA}]} (\delta_{\text{NB}} - \delta_{\text{Free}}) \quad (3)$$

where δ_{Free} is the chemical shift of the free Meis1-HD; and $[\text{MI}]$ and $[(\text{MI})_2 : \text{DNA}]$ are the concentrations of free Meis1-HD and the (MI)₂:DNA complex, respectively, which can be determined as described in the Supplementary Information. The determined ¹H and ¹⁵N chemical shifts of Meis1-HD in the specific and non-specific binding complexes are listed in Supplementary Table 2.

¹⁵N CPMG-RD experiments

The ¹⁵N CPMG-RD experiments were performed using free ¹⁵N-labeled Meis1-HD and ¹⁵N-labeled Meis1-HD complexed with DNA substrates at 298 K⁴⁰. Experiments employed a constant relaxation delay (T_{relax}) of 60 ms and 12 values of $\nu_{\text{CPMG}} = 1/(2\tau_{\text{CP}})$ ranging from 33 to 1000 Hz, where τ_{CP} is the delay between consecutive pulses. Transverse relaxation rates $R_{2,\text{eff}}$ were calculated for each cross-peak signal at each

value by:

$$R_{2,\text{eff}}(\nu_{\text{CPMG}}) = -\frac{1}{T_{\text{relax}}} \ln \left(\frac{I(\nu_{\text{CPMG}})}{I_0} \right) \quad (4)$$

where $I(\nu_{\text{CPMG}})$ and I_0 are the peak intensity at values of ν_{CPMG} of 60 and 0 ms, respectively.

When $[\text{DNA}]_t/[\text{MI}]_t < 0.5$, Meis1-HD shows two conformational exchange processes, a slow exchange of specific binding and a fast exchange of non-specific binding, between the free and complexed forms, which occur independently of each other. In this case, the CPMG-RD data are fitted by⁴¹:

$$R_{2,\text{eff}}(\nu_{\text{CPMG}}) = R_2^0 + \frac{p_{\text{F}} p_{\text{NB}} (\Delta\omega_{\text{FN}})^2}{k_{\text{ex,NB}}} \left\{ 1 - \frac{4\nu_{\text{CPMG}}}{k_{\text{ex,NB}}} \tanh \left(\frac{k_{\text{ex,NB}}}{4\nu_{\text{CPMG}}} \right) \right\} + k_{\text{FS}} \left\{ 1 - \frac{\sin(\Delta\omega_{\text{FS}}/4\nu_{\text{CPMG}})}{\Delta\omega_{\text{FS}}/4\nu_{\text{CPMG}}} \right\} \quad (5)$$

where R_2^0 is the intrinsic transverse relaxation rate; $\Delta\omega_{\text{FS}}$ and $\Delta\omega_{\text{FN}}$ are the chemical shift differences (in Hz) between the free and specific binding states and between the free and non-specific binding states, respectively; p_{F} and p_{NB} are the relative populations of the free and non-specific binding states; k_{FS} is the association rate for the specific DNA binding of Meis1-HD; and $k_{\text{ex,NB}}$ is the exchange rate between free and non-specific binding states of Meis1-HD.

The $k_{\text{ex,NB}}$ was used to calculate the association rate (k_{FN}) using $k_{\text{FN}} = k_{\text{ex,NB}} \times (1 - p_{\text{F}})$ and the dissociation rate (k_{NF}) using $k_{\text{NF}} = k_{\text{ex,NB}} \times p_{\text{F}}$, where p_{F} is the relative population of free Meis1-HD, for non-specific binding. Similarly, the k_{FS} was used to calculate the exchange rate ($k_{\text{ex,SB}}$) using $k_{\text{ex,SB}} = k_{\text{FS}}/(1 - p_{\text{F}})$ and the dissociation rate (k_{SF}) using $k_{\text{SF}} = k_{\text{ex,SB}} \times p_{\text{F}}$ for specific binding. The activation energy difference ($\Delta\Delta G_{\text{off}}^{\ddagger}$) for dissociation of Meis1-HD between the non-specific and specific binding states was calculated using the equation, $\Delta\Delta G_{\text{off}}^{\ddagger} = \Delta G_{\text{off,NB}}^{\ddagger} - \Delta G_{\text{off,SB}}^{\ddagger} = -RT \ln(k_{\text{NF}}/k_{\text{SF}})$.

System preparation for MD simulations

The initial complex structures of Meis1-HD and the meisDNA variants studied in this work were prepared by modifying the crystal structure of the WT complex, PDB ID: 5BNG¹². Starting from the WT crystal structure, the bases of the variants were mutated to the corresponding bases using the LEaP program of the Amber20 molecular simulation package⁴² and the OL15 force field⁴³. All solvated system files were prepared with the CHARMM-GUI webserver⁴⁴⁻⁴⁶. The Meis1-HD structure was parameterized using the ff14SB force field⁴⁷. The initial complex structures were solvated with a cubic TIP3P water box with a dimension of 75.0 Å. K⁺ and Cl⁻ ions were added to make the ionic strength of the solvated systems 0.15 M⁴⁸. Energy minimization calculations were performed for all solvated systems using the steepest descent and conjugated gradient methods for 500 steps. The system details are listed in Supplementary Table 11 and the input coordinates and force field parameter files are provided as Supplementary Data 1.

MD simulations

For each solvated Meis1-HD-DNA system, the energy-minimized initial structures were gradually heated to 300 K for 500 ps with harmonic positional restraints with a force constant of 10 kcal/mol/Å² on all heavy atoms of Meis1-HD and the DNA strands. After the temperature of the system reached 300 K, an equilibrium simulation was performed for 1 ns under constant volume and constant temperature conditions. The temperature of the system was kept constant using the Langevin thermostat⁴⁹. After the equilibrium simulation was finished, a production simulation was performed for 1.5 μs under constant pressure and constant temperature. The Berendsen barostat was used to keep the pressure of the system at 1 atm with a pressure relaxation time of 2 ps⁵⁰. All covalent bonds involving hydrogen were constrained using

the SHAKE algorithm⁵¹, and a timestep of 2 fs was used for all simulations. All MD simulations were performed with the GPU version of the *pmemd* MD engine⁵².

Trajectory analysis of MD simulation results

The last 1 μ s trajectory of each production MD simulation was analyzed using the CPPTRAJ program⁵³. The complex structures were superimposed to the first snapshot of the trajectory, and the root mean square deviation and RMSF were calculated using all heavy atoms. The hydrogen bonds between Meis1-HD and the meisDNA strands were investigated using the *hbond* command. To identify the conformational flexibility of the complexes and identify representative conformations, conformational clustering analysis was performed. The sampled complex conformations were grouped into 10 clusters using the K-means clustering algorithm (Supplementary Data 1). To assess the statistical significance of the differences in RMSF and hydrogen bond patterns between WT and variants, we performed Student's *t*-tests. We estimated the average and standard deviations of the RMSF values using bootstrapping, which involved randomly sampling subsets of the 200 ns trajectories with replacement, conducted 50 times.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data supporting the findings of this study are provided in the Supplementary Information. Source data are provided as a Source Data file. PDB code: [5BNG](#) (TALE type homeobox transcription factor MEIS1 complexes with specific DNA). BMRB code: [52555](#) (Backbone assignments for TALE Homeodomain Transcription Factor Meis1). Source data are provided with this paper.

Code availability

The input coordinates, parameter files, and the centroid structures from the K-means clustering of MD simulations are available as Supplementary Data 1.

References

- Redding, S. & Greene, E. C. How do proteins locate specific targets in DNA? *Chem. Phys. Lett.* **570** <https://doi.org/10.1016/j.cplett.2013.03.035> (2013).
- Pabo, C. O. & Sauer, R. T. Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.* **61**, 1053–1095 (1992).
- Rohs, R. et al. Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.* **79**, 233–269 (2010).
- Privalov, P. L., Dragan, A. I. & Crane-Robinson, C. Interpreting protein/DNA interactions: distinguishing specific from non-specific and electrostatic from non-electrostatic components. *Nucleic Acids Res.* **39**, 2483–2491 (2011).
- Scott, M. P., Tamkun, J. W. & Hartzell, G. W. 3rd The structure and function of the homeodomain. *Biochim. Biophys. Acta* **989**, 25–48 (1989).
- Gehring, W. J. et al. The structure of the homeodomain and its functional implications. *Trends Genet.* **6**, 323–329 (1990).
- Longobardi, E. et al. Biochemistry of the tale transcription factors PREP, MEIS, and PBX in vertebrates. *Dev. Dyn.* **243**, 59–75 (2014).
- Burglin, T. R. & Affolter, M. Homeodomain proteins: an update. *Chromosoma* **125**, 497–521 (2016).
- Blasi, F., Bruckmann, C., Penkov, D. & Dardaei, L. A tale of TALE, PREP1, PBX1, and MEIS1: interconnections and competition in cancer. *Bioessays* **39**, <https://doi.org/10.1002/bies.201600245> (2017).
- Shen, W. F. et al. AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. *Mol. Cell. Biol.* **17**, 6448–6458 (1997).
- Chang, C. P. et al. Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. *Mol. Cell. Biol.* **17**, 5679–5687 (1997).
- Jolma, A. et al. DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* **527**, 384–388 (2015).
- Penkov, D. et al. Analysis of the DNA-binding profile and function of TALE homeoproteins reveals their specialization and specific interactions with Hox genes/proteins. *Cell Rep.* **3**, 1321–1333 (2013).
- Zhang, X., Friedman, A., Heaney, S., Purcell, P. & Maas, R. L. Meis homeoproteins directly regulate Pax6 during vertebrate lens morphogenesis. *Genes Dev.* **16**, 2097–2107 (2002).
- Williams, T. M., Williams, M. E. & Innis, J. W. Range of HOX/TALE superclass associations and protein domain requirements for HOXA13:MEIS interaction. *Dev. Biol.* **277**, 457–471 (2005).
- Shen, W. F. et al. HOXA9 forms triple complexes with PBX2 and MEIS1 in myeloid cells. *Mol. Cell. Biol.* **19**, 3051–3061 (1999).
- Shanmugam, K., Green, N. C., Rambaldi, I., Saragovi, H. U. & Featherstone, M. S. PBX and MEIS as non-DNA-binding partners in trimeric complexes with HOX proteins. *Mol. Cell. Biol.* **19**, 7577–7588 (1999).
- Clamp, M. et al. Distinguishing protein-coding and noncoding genes in the human genome. *Proc. Natl Acad. Sci. USA* **104**, 19428–19433 (2007).
- Kimchi-Sarfaty, C. et al. A “silent” polymorphism in the MDR1 gene changes substrate specificity. *Science* **315**, 525–528 (2007).
- Lu, Y. F. et al. IFNL3 mRNA structure is remodeled by a functional non-coding polymorphism associated with hepatitis C virus clearance. *Sci. Rep.* **5**, 16037 (2015).
- Li, G. et al. Trinucleotide base pair stacking free energy for understanding TF-DNA recognition and the functions of SNPs. *Front. Chem.* **6**, 666 (2018).
- Deplancke, B. et al. A gene-centered *C. elegans* protein-DNA interaction network. *Cell* **125**, 1193–1205 (2006).
- Fuxman Bass, J. I. et al. Human gene-centered transcription factor networks for enhancers and disease variants. *Cell* **161**, 661–673 (2015).
- Qu, X., Cheng, L., Zhao, L., Qiu, L. & Guo, W. Functional variation of SLC52A3 rs13042395 predicts survival of Chinese gastric cancer patients. *J. Cell. Mol. Med.* **24**, 12550–12559 (2020).
- Page, S. H., Wright, E. K. Jr., Gama, L. & Clements, J. E. Regulation of CCL2 expression by an upstream TALE homeodomain protein-binding site that synergizes with the site created by the A-2578G SNP. *PLoS ONE* **6**, e22052 (2011).
- Winter, R. B., Berg, O. G. & von Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The *Escherichia coli* lac repressor-operator interaction: kinetic measurements and conclusions. *Biochemistry* **20**, 6961–6977 (1981).
- Gowers, D. M., Wilson, G. G. & Halford, S. E. Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. *Proc. Natl Acad. Sci. USA* **102**, 15883–15888 (2005).
- Givaty, O. & Levy, Y. Protein sliding along DNA: dynamics and structural characterization. *J. Mol. Biol.* **385**, 1087–1097 (2009).
- Marcovitz, A. & Levy, Y. Obstacles may facilitate and direct DNA search by proteins. *Biophys. J.* **104**, 2042–2050 (2013).
- Ryu, K. S., Tugarinov, V. & Clore, G. M. Probing the rate-limiting step for intramolecular transfer of a transcription factor between specific sites on the same DNA molecule by (15)Nz-exchange NMR spectroscopy. *J. Am. Chem. Soc.* **136**, 14369–14372 (2014).
- Riggs, A. D., Bourgeois, S. & Cohn, M. The lac repressor-operator interaction. 3. Kinetic studies. *J. Mol. Biol.* **53**, 401–417 (1970).

32. Ferretti, E. et al. Segmental expression of Hoxb2 in r4 requires two separate sites that integrate cooperative interactions between Prep1, Pbx and Hox proteins. *Development* **127**, 155–166 (2000).
33. LeBrun, D. P. & Cleary, M. L. Fusion with E2A alters the transcriptional properties of the homeodomain protein PBX1 in t(1;19) leukemias. *Oncogene* **9**, 1641–1647 (1994).
34. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
35. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).
36. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).
37. Hwang, T.-L., Mori, S., Shaka, A. J. & van Zijl, P. C. M. Application of phase-modulated CLEAN Chemical EXchange Spectroscopy (CLEANEX-PM) to detect water–protein proton exchange and intermolecular NOEs. *J. Am. Chem. Soc.* **119**, 6203–6204 (1997).
38. Lee, J. H. & Pardi, A. Thermodynamics and kinetics for base-pair opening in the P1 duplex of the Tetrahymena group I ribozyme. *Nucleic Acids Res.* **35**, 2965–2974 (2007).
39. Lee, J. H., Jucker, F. & Pardi, A. Imino proton exchange rates imply an induced-fit binding mechanism for the VEGF165-targeting aptamer, Macugen. *FEBS Lett.* **582**, 1835–1839 (2008).
40. Palmer, A. G. 3rd, Kroenke, C. D. & Loria, J. P. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol.* **339**, 204–238 (2001).
41. Lee, A. R. et al. NMR dynamics study reveals the α domain of human ADAR1 associates with and dissociates from Z-RNA more slowly than Z-DNA. *ACS Chem. Biol.* **14**, 245–255 (2019).
42. Case, D. A. et al. The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).
43. Galindo-Murillo, R. et al. Assessing the current state of Amber force field modifications for DNA. *J. Chem. Theory Comput.* **12**, 4114–4127 (2016).
44. Lee, J. et al. CHARMM-GUI supports the Amber force fields. *J. Chem. Phys.* **153**, <https://doi.org/10.1063/5.0012280> (2020).
45. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, <https://doi.org/10.1002/jcc.20945> (2008).
46. Brooks, B. R. et al. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**, <https://doi.org/10.1002/jcc.21287> (2009).
47. Maier, J. A. et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
48. Joung, I. S. & Cheatham, T. E. 3rd Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).
49. Loncharich, R. J., Brooks, B. R. & Pastor, R. W. Langevin dynamics of peptides: the frictional dependence of isomerization rates of N-acetylalanyl-N'-methylamide. *Biopolymers* **32**, 523–535 (1992).
50. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3684 (1984).
51. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
52. Salomon-Ferrer, R., Gotz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J. Chem. Theory Comput.* **9**, 3878–3888 (2013).
53. Roe, D. R. & Cheatham, T. E. 3rd PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).

Acknowledgements

This work was supported by the Samsung Science and Technology Foundation (SSTF-BA1701-10 to S.-R.C., Y.-J.S., H.-S.J., H.-B.A., Y.G., and J.-H.L.) and the National Research Foundation of Korea [Grants 2020R1A2C1006909 and 2022R1A4A1021817 (BRL) to S.-R.C., Y.-J.S., H.-S.J., H.-B.A., Y.G., and J.-H.L., 2020M3A9G7103933, 2022M3E5F3081268, and 2022R1C1C1005080 to J.L.]. This work was supported by the Institutional Research Program of Korea Institute of Science and Technology (KIST; 2V10021 to S.-R.C. and N.-K.K.). This research was supported by the BK21FOUR Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Education (5120200513755 to J.L.). We thank the KBSI Protein Structure Team and KIST Advanced Analysis Data Center for the NMR experiments and Melissa Stauffer, of Scientific Editing Solutions, for editing the manuscript.

Author contributions

J.-H.L. conceived and designed the experiments; S.-R.C. and Y.-J.S. performed sample preparation and ITC analysis; S.-R.C., H.-S.J., H.-B.A., Y.G., N.-K.K., and K.-S.R. performed NMR experiments; J.L. performed MD simulation and analysis; S.-R.C., J.L., and J.-H.L. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-51297-7>.

Correspondence and requests for materials should be addressed to Juyong Lee.

Peer review information *Nature Communications* thanks Francesco Blasi and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024