BBA - Proteins and Proteomics 1866 (2018) 1021–1028

Contents lists available at ScienceDirect

BBA - Proteins and Proteomics

journal homepage: www.elsevier.com/locate/bbapap

Characterization of the interaction interface and conformational dynamics of human TGIF1 homeodomain upon the binding of consensus DNA

Shuangli Li[a,b], Rui Hu[a], Haijie Yao[c], Dong Long[c,d], Fan Luo[a], Xin Zhou[a], Xu Zhang[a], Maili Liu[a], Jiang Zhu[a,*], Yunhuang Yang[a,*]

[a] State Key Laboratory of Magnetic Resonance and Atomic Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Chinese Academy of sciences, Wuhan 430071, China
[b] Graduate University of Chinese Academy of Sciences, Beijing 100049, China
[c] School of Life Sciences, University of Science and Technology of China, Hefei 230027, China
[d] Hefei National Laboratory for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei 230027, China

ARTICLE INFO

Keywords:
TG interacting factor-1
DNA recognition
Interaction interface
Conformational dynamics
NMR spectroscopy

ABSTRACT

The TG interacting factor-1 homeodomain (TGIF1-HD) binds with the consensus DNA motif 5′-TGTCA-3′ in gene promoters through its three-amino acid loop extension (TALE) type homeodomain, and then recruits co-regulators to regulate gene expression. Although the solution NMR structure of human TGIF1-HD has been reported previously, little is known about its DNA binding mechanism. NMR titrations have been extensively used to study mechanisms of ligand binding to target proteins; however, an intermediate exchange occurred predominantly between TGIF1-HD in the free and bound states when titrated with the consensus DNA, which resulted in poor-quality NMR spectra and precluded further exploration of its interaction interface and conformational dynamics. Here, the helix α3 of TGIF1-HD was speculated as the specific DNA binding interface by hydrogen–deuterium exchange mass spectrometry (HDX-MS) experiments, and subsequently confirmed by chemical exchange saturation transfer (CEST) spectroscopy. In addition, simultaneous conformational changes in other regions, including α1 and α2, were induced by DNA binding, explaining the observation of chemical shift perturbations from extensive residues besides those located in α3. Further, low-populated DNA-bound TGIF1-HD undergoing a slow exchange at a rate of 130.2 ± 3.6 s^{-1} was derived from the analysis of the CEST data, and two residues, R220 and R221, located in the middle of α3 were identified to be crucial for DNA binding. Our study provides structural and dynamic insights into the mechanisms of TGIF1-HD recognition of extensive promoter DNA.

1. Introduction

The TG interacting factor-1 (TGIF1) protein in humans controls fetal brain development, as well as the balance between cell proliferation and differentiation in adult tissues [1–3]. To date, TGIF1 has been reported to be involved in many human diseases. Loss of TGIF1 function causes holoprosencephaly (HPE), a severe human genetic disease with abnormal craniofacial development [3–7]. TGIF1 is also implicated in many kinds of cancers, such as acute myelogenous leukemia (AML), hepatocellular carcinoma (HCC), urothelial carcinoma (UC), and triple-negative breast cancer (TNBC) [2, 8–10]. Human TGIF1, which functions mainly as a transcriptional repressor [11, 12], contains 401 amino acids and consists of three functional domains, the homeodomain (HD) and repressive domains 1 and 2 (RD1 and RD2) (Fig. 1A). Extensive studies have revealed that HD functions in specific DNA binding, while RD1 and RD2 recruit the transcriptional co-repressors carboxyl terminus-binding protein (CTBP) and Sin3A, respectively [1, 13–15].

In particular, HDs generally comprise three α-helices and are engaged by numerous transcription factors (TFs) for binding sequence-specific DNA, which is the initial step in TF-mediated transcription regulation [16]. The HD domain of TGIF1 (TGIF1-HD; Fig. 1A) belongs to the three-amino acid loop extension (TALE) superclass of HDs, which contain three more amino acids in the loop connecting α-helices 1 and 2 [17]. TGIF1-HD specifically binds a usual consensus DNA motif, 5′-TGTCA-3′ (Fig. 1B), which harbors G and C substitutions in the TA-rich motifs bound by canonical HD [1]. Deletion of HD and missense

* Corresponding authors at: State Key Laboratory of Magnetic Resonance and Atomic Molecular Physics, Wuhan Center for Magnetic Resonance, Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan 430071, China.
E-mail addresses: jiangzhu@wipm.ac.cn (J. Zhu), yang_yh@wipm.ac.cn (Y. Yang).

https://doi.org/10.1016/j.bbapap.2018.07.005
Received 26 March 2018; Received in revised form 28 May 2018; Accepted 17 July 2018
Available online 23 July 2018
1570-9639/

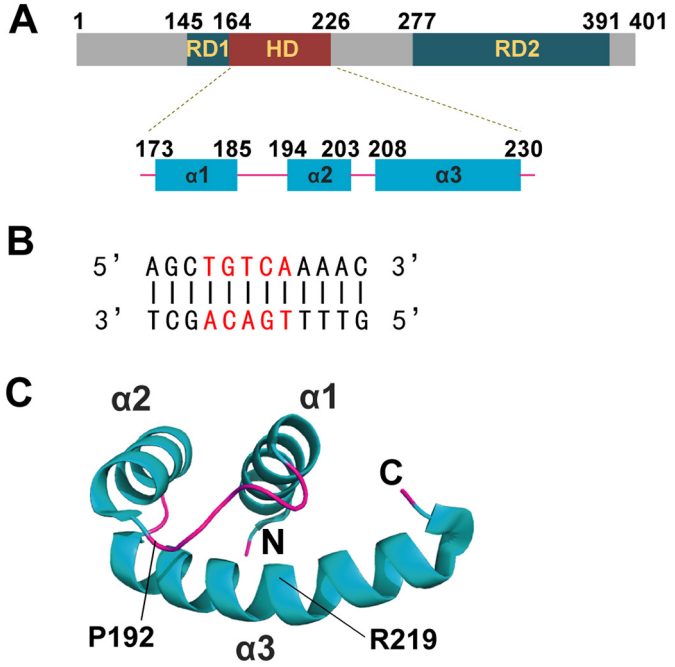

Fig. 1. Schematic diagram of human TG interacting factor-1 (TGIF1) and solution structure of TGIF1 homeodomain (TGIF1-HD). (A) Full-length human TGIF1 protein contains 401 amino acids. Three defined functional domains are homeodomain (HD) and repressive domains 1 and 2 (RD1 and RD2) (top panel). The HD consists of three α-helices (bottom panel). The numbers of beginning and ending amino acids for each domain and helix are shown. (B) The double-stranded DNA used in this study. The consensus motif specifically bound by TGIF1-HD is colored in red. (C) The solution NMR structure of TGIF1-HD (PDB ID: 2LK2). Two holoprosencephaly (HPE)-related residues are indicated.

mutations in HD that result in loss of DNA-binding activity abolish the transcriptional regulation of TGIF1 and cause HPE disease [4, 11]. A genome-scale study in myeloid leukemia cells revealed that TGIF1 could directly bind the 5′-TGTCA-3′ motifs in promoters of > 1600 genes and regulate their transcription to modulate the proliferation and differentiation of myeloid leukemia cells [2]. Subsequent studies identified *SOX3* and *sterol O-acyltransferase 2* (*SOAT2*) as the target genes of TGIF1-mediated transcriptional regulation during brain development and cholesterol metabolism, respectively [18, 19].

To better understand the biological functions of TGIF1, we focused on the study of its structure and mechanism of molecular interactions. In our previous study, we reported the solution NMR structure of TGIF1-HD (Fig. 1C), and two HPE-related mutations were confirmed to affect TGIF1-HD structure and DNA binding affinity. In addition, isothermal titration calorimetry (ITC) experiments were carried out to obtain the consensus DNA (Fig. 1B) binding affinity of TGIF1-HD with a dissociation constant of 0.26 μM and a binding stoichiometry of 1 [7]. At present, the structure of TGIF1-HD in complex with consensus DNA has not yet been determined. An intermediate exchange occurring on TGIF1—HD binding to DNA led to most NMR peaks being undetectable due to line broadening in the 2D $^{1}H—^{15}N$ heteronuclear single quantum coherence (HSQC) spectrum, severely hindering investigations of the DNA recognition mechanism. In this study, we determined the middle of alpha helix 3 (α3) as the specific DNA binding site based on analyses of hydrogen–deuterium exchange mass spectrometry (HDX-MS) experiments and chemical exchange saturation transfer (CEST) spectroscopy. Besides, simultaneous conformational changes in other regions were induced by DNA binding. Despite the predominant intermediate conformational exchange, a few DNA-bound TGIF1-HDs in a slow exchange regime at the millisecond level were detected by CEST spectroscopy. Also, two residues, R220 and R221, located in the highly conserved helix 3, were determined as key residues for DNA binding.

2. Materials and methods

2.1. Preparations of protein and dsDNA samples

TGIF1-HD and double-stranded DNA (5′-AGCTGTCAAAAC-3′) samples were prepared as described previously [7]. R220A and R221A versions of TGIF1-HD were expressed and purified following the methods for wild-type TGIF1-HD.

2.2. NMR titration experiments and chemical shift assignments for holo-TGIF1-HD

NMR titration experiments were performed on the Bruker 600 MHz instrument at 298 K. The samples containing 0.40 mM TGIF1-HD and different concentrations of DNA (0 mM, 0.20 mM, 0.40 mM, and 0.80 mM) in the NMR buffer containing 10% (v/v) D_2O, 20 mM NH_4OAc, 100 mM NaCl, 5 mM $CaCl_2$, 10 mM DTT, and 0.02% NaN_3, at pH 4.5, were premixed and allowed to equilibrate for over 1 h. Chemical shift assignment of holo-TGIF1-HD was carried out using a sample containing 30 μL of 5 mM DNA and 270 μL of 0.4 mM *U*-[^{13}C, ^{15}N]-labeled samples of TGIF1-HD with 10% D_2O. Based on NMR data including HNCO, HNCA, HNCOCA, HNCACB, and CBCACONH, backbone resonances were automatically assigned first using the PINE server [20], and then confirmed manually.

2.3. Hydrogen–deuterium exchange mass spectrometry (HDX-MS)

Peptide-level HDX-MS was performed as follows. Both apo- and holo-protein samples were diluted with equilibrium buffer (100 mM sodium phosphate, pH 6.8) to a concentration of 40 μM. Briefly, at the start of an HDX reaction, at time zero ($t = 0$), protein solution at 40 μM was constituted with labeling buffer (100 mM sodium phosphate, D_2O, pD 6.4) at an approximate 13-fold dilution. The labeling mixtures were incubated at 20 °C, and deuterium exchange was quenched at the following time points: 10 s, 1 min, 10 min, and 60 min. Deuterium labelling was quenched by making a 1:1 dilution with chilled quenching buffer (100 mM sodium phosphate, pH 2.3). Quenched samples were digested, desalted, and separated online using an Acquity UPLC M-Class system with HDX-2 automation coupled with Synapt G2-Si HDMS. The online digestion was performed using an immobilized pepsin column, 2.1 mm × 30 mm (Enzymate Pepsin Column, Waters Corp., Milford, MA, USA) for 4 min in 0.1% formic acid, with H_2O at a flow rate of 70 μL/min. The entire digestion was held at 15 °C within the column compartment of the HDX manager. The peptides were then trapped and desalted online, and then separated with an Acquity UPLC BEH C18 1.7 μm, 1 mm × 100 mm column (Waters) held at 0 °C. After 6 min of linear elution, eluent was directed into a Synapt G2-Si HDMS and lockmass corrected. Mass spectra were acquired in MS^E mode over the mass range of 100 to 2000. Blank injections were inserted after each sample injection to eliminate the effect of protein carryover. The data was then processed with ProteinLynx Global Server (Waters) for peptide sequence coverage and DynamX 3.0 (Waters) to automatically calculate deuterium uptake numbers at different time points and generate the differential graph along the sequence.

2.4. Chemical exchange saturation transfer (CEST) spectroscopy

All ^{15}N CEST experiments [21] were performed at 298 K on a Varian Inova 700 MHz spectrometer equipped with a cryogenic probe. The exchange duration (T_{EX}) was 400 ms. A pair of weak B_1 field strengths (12 Hz and 26 Hz) were used, calibrated using the method of Guenneugues and Berthault [22]. A series of 2D $^{1}H—^{15}N$ CEST spectra were recorded with the frequency of the weak B_1 field ranging from 100 to 133 ppm in a step size of 20 Hz, along with a reference spectrum

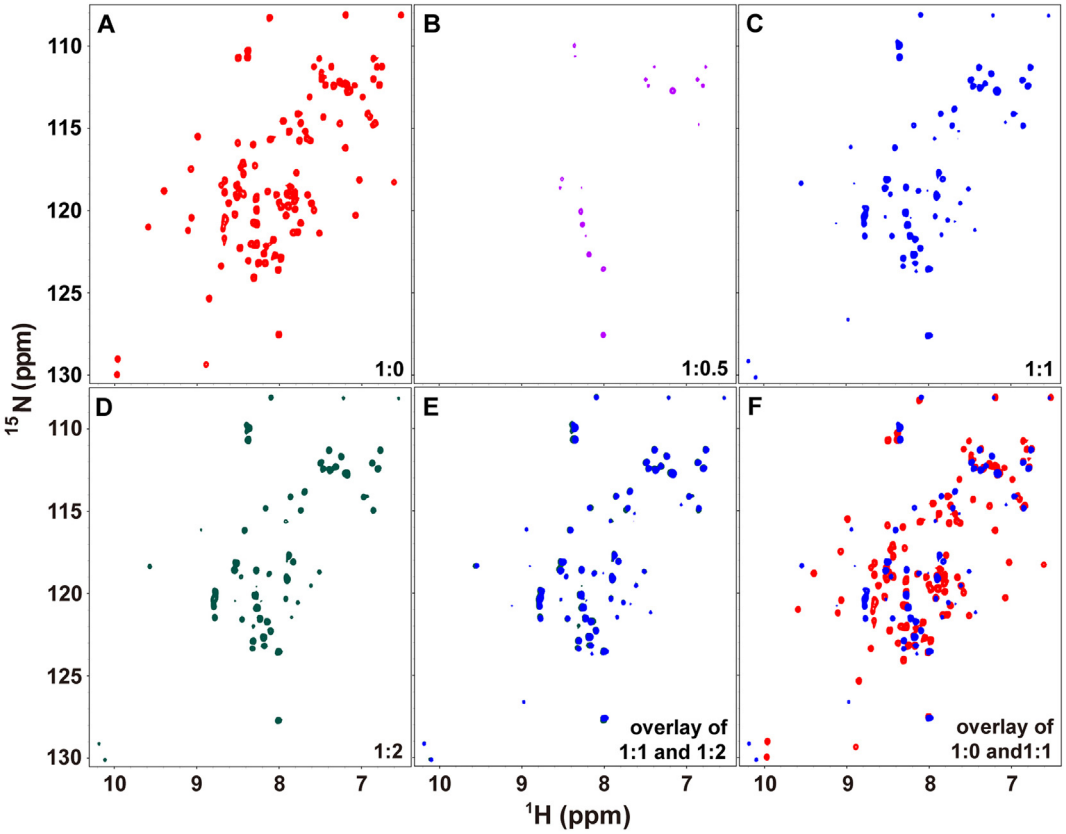

Fig. 2. 2D ^{1}H−^{15}N heteronuclear single quantum coherence (HSQC) spectra for NMR titrations of TGIF1-HD with consensus DNA at a series of protein-to-DNA molar ratios as indicated: (A) 1:0; (B) 1:0.5; (C) 1:1; (D) 1:2. (E) Overlay of (C) and (D); (F) overlay of (A) and (C).

(T_{EX} = 0). The spectra were acquired with 2 scans/FID and 100 complex points for the indirect dimension, with the total measurement time of 32 h (for each B_1 field). All CEST spectra were processed using NMRPipe and analyzed with the ChemEx program (https://github.com/gbouvignies/chemex). We used a Monte Carlo analysis to obtain uncertainties of the fitted parameters of the exchange rate (k_{ex}) and population of DNA-bound complex in the excited state (p_B).

2.5. Circular dichroism (CD) spectra

CD spectra were recorded on a Chirascan™ CD spectrometer (Applied Photophysics) using 0.2 cm path-length quartz cells, following the method described previously [7].

2.6. Isothermal titration calorimetry (ITC) measurements

ITC experiments were performed on a VP-ITC instrument (GE Healthcare). All samples were prepared in the same buffer as used in the NMR experiments to avoid any heat change resulting from mixing buffers. The protein solutions of wild-type, R220A, and R221A versions of TGIF1-HD in 100 μM were titrated at 298 K into the reservoir containing 10 μM DNA. The experiments and data analysis were carried out as described previously [7]. All ITC experiments were repeated in triplicate.

3. Results and discussion

3.1. A typical intermediate exchange occurred between TGIF1-HD in free and DNA-bound states

NMR titration experiments are very powerful for studying mechanisms of ligand binding to target proteins through screening NMR observables, including chemical shift and cross-peak intensity. The 12 bp dsDNA (5′-AGCTGTCAAAAC-3′), as mentioned earlier (Fig. 1B), was selected for titration to characterize the interaction interface and conformational dynamics of TGIF1-HD. A series of 2D ^{1}H−^{15}N HSQC spectra were recorded for mixtures of TGIF1-HD:DNA at different molar ratios, as indicated in Fig. 2. It can be clearly seen that most of the protein NMR resonances disappeared upon the addition of DNA due to line broadening (Fig. 2B), and reappeared to some extent when the stoichiometry of the complex was reached (Fig. 2C), which usually indicates a typical intermediate exchange taking place on the chemical-shift time scale between proteins in free and bound states.

The spectrum remained unchanged when further titrated with excess DNA (Fig. 2D and E), suggesting that the DNA binding was saturated and the stoichiometry for the protein-DNA complex was 1. Estimated approximately, about 60 cross-peaks were present for the complex sample, coming from amino acid backbone and side-chain NH, as well NH_2 of Asn and Gln side-chains. Resonances observed were much less than for protein alone, and only 28 of these resonances could be unambiguously assigned based on NMR data collected for backbone chemical shift assignment using the *U*-[^{13}C/^{15}N]-labeled TGIF1-HD:DNA complex (Table S1). The spectrum of TGIF1-HD in the complex became poor due to conformational dynamics on an unfavorable NMR time scale, similar to some other nucleic acid binding proteins, as previously reported [23–25]. Therefore, poor spectrum for the complex with insufficient resonance and weak NMR signal intensity severely hindered further investigation of the binding interface and interaction mechanism using NMR spectroscopy. It was also noted that NMR signal intensities were not uniform, indicating the existence of heterogeneous conformations in the system under study. Overlaying the spectra for apo- and holo-proteins (Fig. 2F), chemical shift perturbations (CSPs) were surprisingly observed for a lot of resonances, far more than expected for those only from residues located in or close to the specific

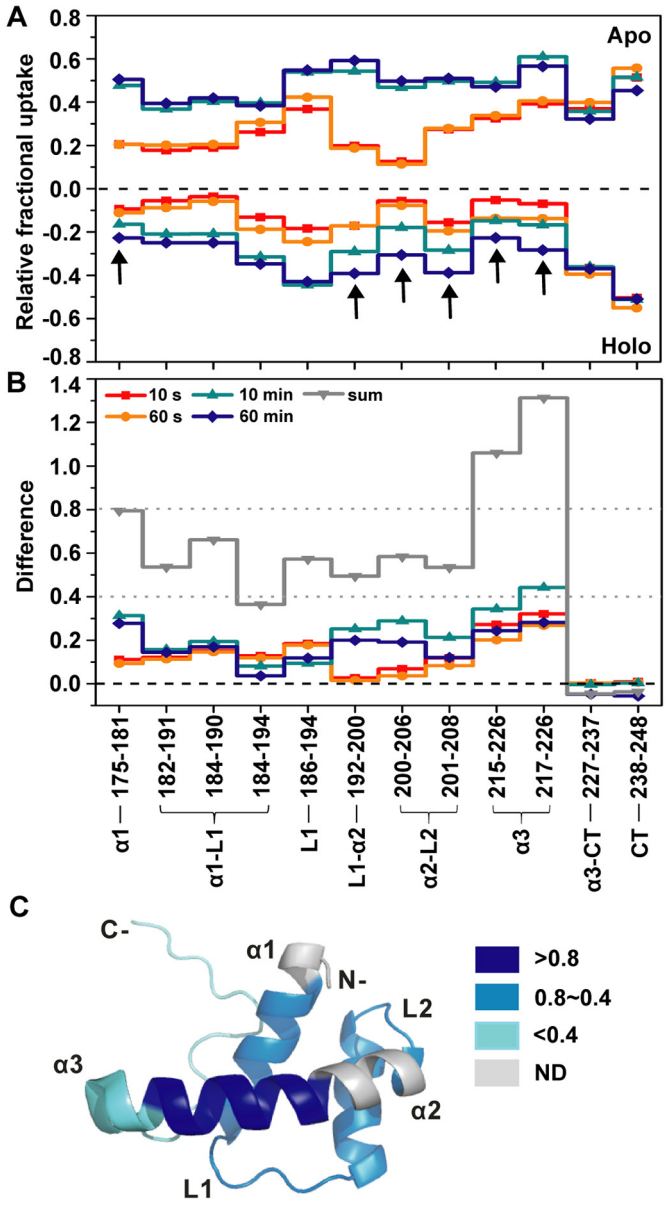

Fig. 3. The DNA binding interface and conformational change of TGIF1-HD revealed by hydrogen–deuterium exchange mass spectrometry (HDX-MS). (A) A mirror plot comparing the relative deuterium exchange for each peptide detected from the N to C terminus of apo- and holo-TGIF1-HD. The relative deuterium uptake for each peptide determined at 10 s, 1 min, 10 min, and 60 min time points is shown. The segments of holo-TGIF1-HD with relative deuterium uptake at 60 min greater than that at 10 min are indicated by arrows. (B) A difference plot, plotting the difference in relative deuterium uptake between apo- and holo-TGIF1-HD. The difference of relative deuterium uptake for each peptide determined at 10 s, 1 min, 10 min, and 60 min time points, as well as the sum of four time point differences (gray) are shown. CT, C-terminus. (C) The sum of four time point differences for each segment in (B) is mapped onto the structure of TGIF1-HD, with color labeling as indicated. ND, not determined.

DNA binding site, suggesting that conformational changes at other regions also took place due to the binding of DNA.

3.2. DNA binding interface of TGIF1-HD characterized by HDX-MS

The DNA binding interface on the target protein can be determined by using HDX-MS to detect the trend of the deuterium exchange rate of target proteins by measuring the molecular mass of online digested peptides at different time points [26]. To do this, apo- and holo-TGIF1-HD were each mixed at low concentrations into bulk D_2O for H–D exchange for either 10 s, 1 min, 10 min, or 60 min, and then cleaved by pepsin for mass determination. As shown in Fig. 3A, the mirror plot demonstrates the relative fractional deuterium uptake of 12 selected protein segments covering the whole protein chain. In general, the relative deuterium uptake became greater as the H–D exchange time increased until H–D exchange was completed. For apo-TGIF1-HD, the pattern of relative deuterium uptake of most segments at 10 s was almost the same as that at 60 s. However, higher relative deuterium uptake was apparently found at either 10 min or 60 min for all segments, except for segments from the C-terminus (Fig. 3A, top panel). By comparison, the relative deuterium uptake at 60 min was almost identical to that at 10 min, suggesting that complete H–D exchange for each segment was nearly reached within 10 min. In contrast, this was not the case for the holo-TGIF1-HD. In particular, the relative deuterium uptake for the segments kept increasing even at 60 min (Fig. 3A, bottom panel, indicated by arrows), suggesting slower H–D exchange taking place in these regions.

In order to highlight the changes in H–D exchange between apo- and holo-proteins, the difference in relative deuterium uptake was plotted for each selected segment at each H–D exchange time point (Fig. 3B), as well the sum of the relative deuterium uptake from all four time points. Obviously, two C-terminal segments, 227–237 and 238–248, exhibited nearly no difference in relative deuterium uptake between apo- and holo- samples, accounting for no direct DNA binding at these regions. In contrast, segments 215–226 and 217–226, located in α3, were observed to have the most significant differences in relative deuterium uptake (> 0.8). We speculate here that helix 3 of TGIF1-HD is the specific DNA binding interface, also inferred from reports that helix 3 of some other TGIF1-HD homologs were confirmed as the specific DNA binding site [27, 28]. In addition, less difference in relative deuterium uptake (around 0.4–0.8) was derived for the rest of the segments covering residues from S175 to T208 (Fig. 3B), which might be reasonably explained by the formation of relatively compact conformation of these regions induced by the DNA binding. All segments with different relative deuterium uptake were mapped to the apo-NMR structure of TGIF1-HD to illustrate the overall conformational changes upon DNA binding (Fig. 3C). The observed conformational changes in other regions including α1 and α2 were consistent with our previous findings of chemical shift perturbations coming from extensive residues in the NMR titrations.

3.3. Low populated DNA-bound TGIF1-HD in slow exchange regime at millisecond level derived by CEST spectroscopy

Taking the protein-RNA complex as an example, the dissociation constant of protein complex in the intermediate exchange regime was reported between 400 and 2 μM [29]. In general, protein complexes characteristically have high affinity (lower dissociation constant) in a slow exchange regime, and relatively low affinity (higher dissociation constant) in a fast exchange regime. The dissociation constant of 0.26 μM for the DNA-bound TGIF1-HD complex fitted from the ITC data is a little lower than expected for a typical intermediate exchange process. We speculate that our system might contain low populated DNA-bound complex in the slow exchange regime with relatively high binding affinity. CPMG relaxation dispersion experiments and CEST spectroscopy are two recently developed NMR methods that are reliable for studying protein conformational dynamics in slow exchange ranging from microsecond to millisecond and detecting invisible low-populated states of proteins in excited states that are usually essential for functioning [21, 30–32].

The two methods mentioned above were employed to test whether slow exchanges could take place in the system under study. Out of consideration for better NMR spectrum quality, we made the mixed sample containing 0.7 mM TGIF1-HD and 0.07 mM DNA with a protein-

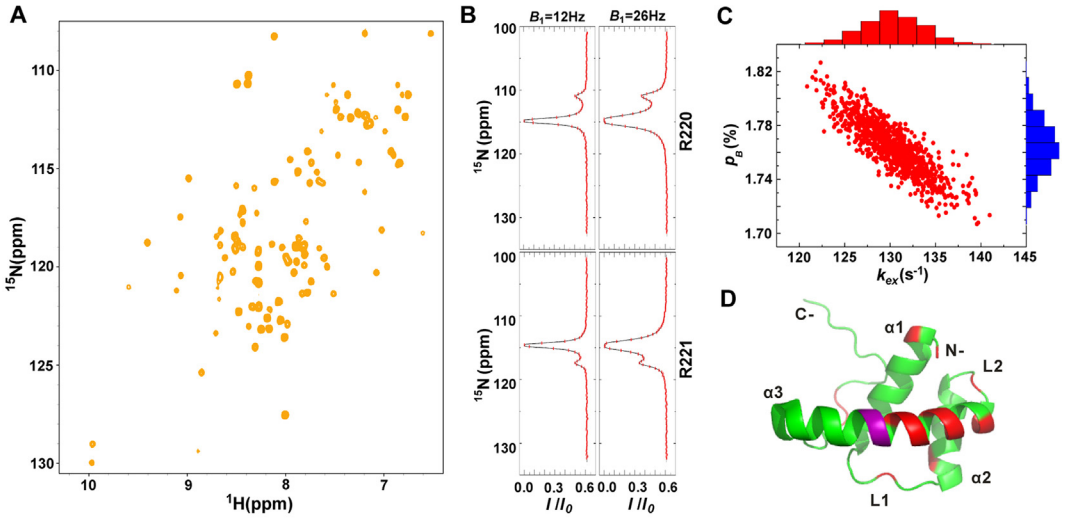

Fig. 4. Low populated TGIF1-HD complex undergoes a slow conformational exchange. (A) 2D ^{1}H−^{15}N HSQC spectrum of TGIF1-HD:DNA at the molar ratio of 1:0.1. (B) Representative CEST profiles for R220 and R221, derived from datasets measured using two kinds of B_1 field strengths as indicated. The solid curves are the best fits with a two-state model (free ↔ bound). (C) Conformational exchange parameters (k_{ex} and p_B) of TGIF1-HD obtained from a global fit of all CEST profiles. Distributions of k_{ex} and p_B were obtained by Monte Carlo analysis. (D) The DNA binding interface of TGIF1-HD drawn by CEST. The residues with $^{\triangle}\omega_{CEST} > 0.5$ are mapped onto the structure of TGIF1-HD and shown in red. R221 and R221 are labeled with purple.

Table 1
Residues with the fitted $|^{\triangle}\omega_{CEST}|$ values > 0.5 ppm derived from chemical exchange saturation transfer (CEST) experiments.

Secondary structure	Residue number	$^{\triangle}\omega_{CEST}$ (ppm)	Standard deviation (ppm)
NT-	171	1.57	0.008
α1	173	1.12	0.008
	190	−0.51	0.184
α2	196	0.54	0.016
L2	206	0.58	0.014
α3	208	0.66	0.012
	209	0.84	0.013
	212	1.10	0.101
	213	1.18	0.039
	214	0.56	0.016
	216	1.04	0.013
	217	1.31	0.012
	218	1.36	0.010
	220	−3.94	0.006
	221	2.86	0.007
CT-	237	−0.86	0.006
	238	−0.87	0.005

to-DNA molar ratio of 1:0.1 (Fig. 4A) instead of 1:0.5 (Fig. 2B). On one side, the fitted CPMG relaxation dispersion curves revealed no conformational exchange for any resonance at microsecond level (data not shown). On the other side, CEST data fittings using a global two-state exchange model demonstrated that two minor dips were clearly observed for residues R220 and R221 in the amide ^{15}N intensity profiles (Fig. 4B), and the values of chemical shift difference, $^{\triangle}\omega_{CEST}$ ($\delta_{minor} - \delta_{major}$), were −3.94 and 2.86 ppm for R220 and R221, respectively. The CEST experiments were performed at two different saturating fields, 12 and 26 Hz. As expected, the resolution for the two fitted resonances improved as the saturating field was decreased. After a detailed analysis of CEST data, 17 residues with $^{\triangle}\omega_{CEST} > 0.50$ ppm were derived (Table 1). Of those, 10 residues were located in helix 3, whereas 7 residues were located in other regions.

At the same time, $^{\triangle}\omega_{CEST}$ values obtained from fittings of CEST profiles were in reasonable agreement with directly measured chemical shift differences between apo- and holo-TGIF1-HD, $^{\triangle}\omega_{direct}$ ($\delta_{holo-} - \delta_{apo-}$), at least for the assigned 28 residues in the complex, as mentioned above (Table S1 and Fig. S1). The exchange rate (k_{ex}) value of 130.2 ± 3.6 s^{-1} and population of DNA-bound TGIF1-HD undergoing this particular kind of slow exchange (p_B) of 1.76 ± 0.02% were further derived from the global fitting of CEST data using all residues with $^{\triangle}\omega_{CEST} > 1.00$ ppm (Fig. 4C). As $^{\triangle}\omega_{CEST}$ and $^{\triangle}\omega_{direct}$ values for at least 28 residues showed reasonable consistency, the low populated holo-TGIF1-HDs in the slow exchange regime were considered to be those binding the DNA in a specific manner, which accounted for the relatively low dissociation constant obtained from ITC experiments.

With respect to the TGIF1-HD-to-DNA molar ratio of 1:0.1 in the CEST sample, the population of the DNA-bound state was calculated to be around 9.4%, based on the protein concentration and the K_D measured by ITC. Considering that TGIF1-HD proteins are largely excess, the unexpected low population of specific DNA-bound state of TGIF1-HD might be the result of a great number of TGIF1-HDs nonspecifically binding with the exposed regions of the DNA in the specific complex. Even more, two TGIF1-HD proteins may nonspecifically bind one DNA molecule simultaneously. In either case, there would be interference with the specific DNA binding of TGIF1-HD. Nevertheless, the k_{ex} value in the slow exchange regime derived from CEST data is consistent with the affinity of TGIF1-HD toward the specific DNA fitted from ITC data, and therefore could represent the exchange rate between free and specific DNA-bound state of TGIF1-HD.

Further, the specific DNA-binding interface of TGIF1-HD could be deduced based on residues with $^{\triangle}\omega_{CEST}$ values > 0.50 ppm. When mapped to the tertiary structure of TGIF1-HD, as shown in Fig. 4D, most of these residues, including C212, N213, W214, I216, N217, A218, R220, and R221, were mainly located in the middle of α3, indicating that this region of TGIF1-HD is for specific DNA binding. This result was consistent with our previous speculation from HDX-MS experiments. In addition, residues R220 and R221 were indicated to be important for the binding of consensus DNA.

3.4. R220 and R221 in helix 3 are key residues for the DNA binding of TGIF1-HD

The structure of apo-TGIF1-HD revealed that the side-chains of P192 and R219 were inward and involved in TGIF1-HD structure packing [7]; by contrast, the side-chains of R220 and R221 were outward and not likely to be involved in TGIF1-HD folding, but might play a role in DNA binding instead. In order to determine whether R220 and

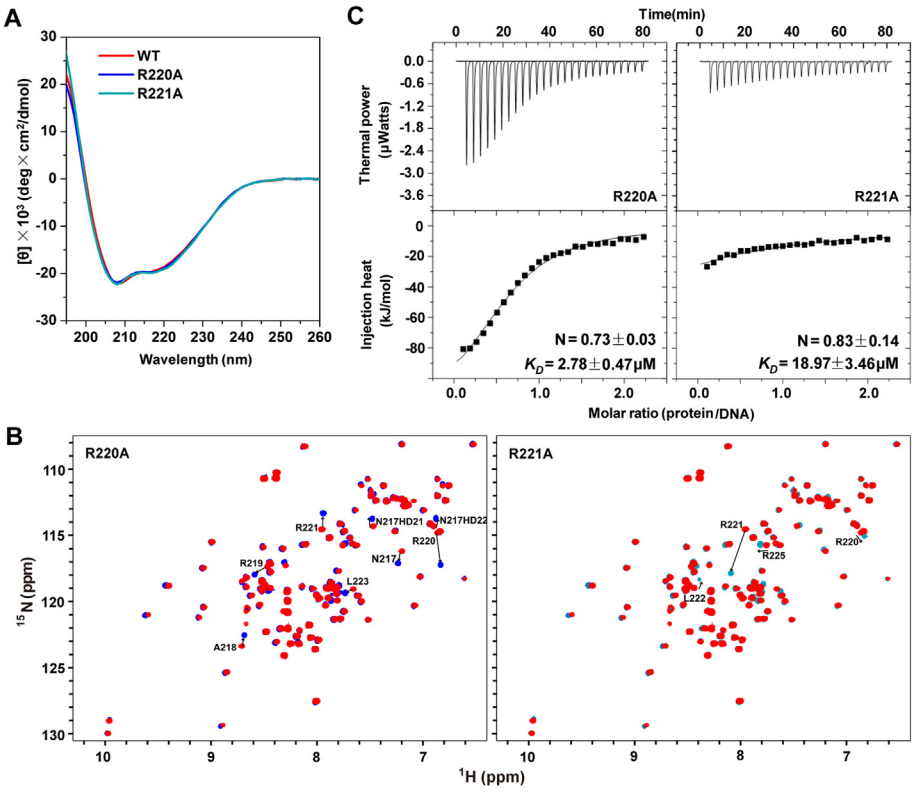

Fig. 5. The roles of R220 and R221 in DNA binding of TGIF1-HD. (A) The far-UV CD spectra of wild-type (red), R220A (blue), and R221A (cyan) versions of TGIF1-HD. (B) Overlaid 2D $^{1}H-^{15}N$ HSQC spectra of wild-type (red) and R220A (blue, left) or R221A (cyan, right) versions of TGIF1-HD. The residues with marked chemical shift perturbations (CSPs) are labeled. (C) The binding of R220A and R221A versions of TGIF1-HD to the DNA containing consensus motif assayed using isothermal titration calorimetry (ITC) experiments. The observed heat changes resulting from injections of protein into DNA are shown in the top panels, while the binding enthalpies evaluated with the assumption of a one-site binding model are shown in the bottom panels. The resulting stoichiometric ratio (N) and dissociation constant (K_D) derived from the fitting are indicated.

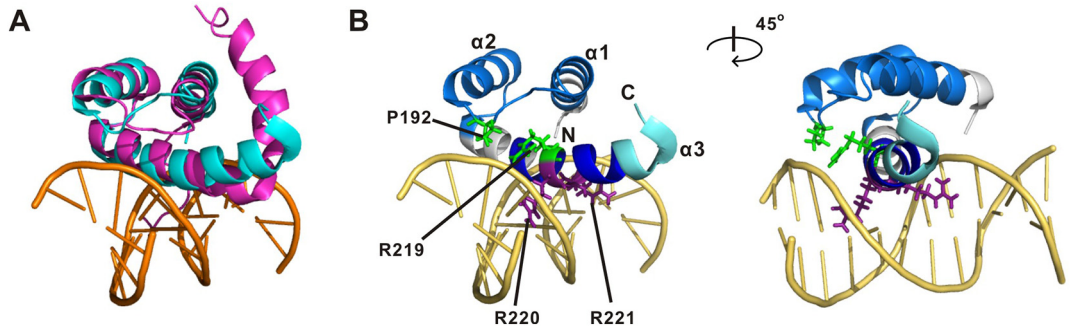

Fig. 6. The DNA binding model of TGIF1-HD derived from HDX-MS and CEST. (A) A structure alignment of free TGIF1-HD (cyan, PDB ID: 2LK2) and PBX1-HD:DNA complex (magenta, PDB ID: 1LFU) with an overall root-mean-square deviation (RMSD) of 1.801 Å. (B) A model of the TGIF1-HD:DNA complex is drawn based on the complex structure of PBX1-HD:DNA (PDB ID: 1LFU). PBX1-HD was substituted by TGIF1-HD in the complex after these two proteins were structurally aligned in (A). Colors used are the same as those in Fig. 3C, except that P192 and R219 are in green, while R220 and R221 are in purple.

R221 are key residues for DNA binding, we designed two mutants, R220A and R221A, that changed the long charged side-chains to short hydrophobic side-chains, respectively. Consistent with this speculation, CD spectra of R220A and R221A mutants were almost identical to that of wild-type (Fig. 5A). Also, $^{1}H-^{15}N$ HSQC spectra of R220A and R221A mutants were highly similar to that of wild-type, and little differences in chemical shift were observed from those residues adjacent to the mutation sites (Fig. 5B), suggesting that these two mutations would not affect the overall structure of TGIF1-HD. Subsequently, the DNA-binding roles of R220 and R221 were tested with these two mutants by ITC experiments. The DNA-binding affinity of R220A and R221A was reduced by 10.7-fold (K_D = 2.78 μM) and 73-fold (K_D = 18.97 μM), respectively (Fig. 5C), which confirmed that R220 and R221 are key residues for DNA binding. By comparison, I289 and R290 are two residues in the corresponding positions in the TALE-type HD of PBX1, a homologous protein of TGIF1, and they make van der Waals interaction or hydrogen bond contacts with the nucleotide bases in the major groove of DNA based on the known structures of PBX1-HD:DNA [27, 33], which also supports the key role of R220 and R221 in DNA binding.

3.5. DNA binding model of TGIF1-HD

Taking our experimental results together, we proposed a model of the complex structure of DNA-bound TGIF1-HD combining the reported complex structure of PBX1-DNA (PDB ID:1LFU) [33]. Basically, protein PBX1 was substituted by TGIF1-HD in the complex after these two proteins were structurally aligned (Fig. 6A). In the TGIF1-HD complex model (Fig. 6B), the middle of helix α3 that was deduced to be a DNA-contacting region of TGIF1-HD by HDX-MS was localized close to the DNA duplex, especially with the side-chains of R220 and R221 stretching into the major groove of the DNA. The regions with medium levels of difference in relative deuterium uptake, including helices α1 and α2, seemed far from the DNA duplex, consistent with the speculation that these regions did not directly contact DNA binding, but might undergo conformational changes upon DNA binding. In contrast,

conformational changes between free and DNA-bound PBX1-HD are mainly located in the α3 helix (Fig. S2).

As mentioned above, I289 and R290, the counterparts of TGIF1 R220 and R221 in PBX1, play an important role in specific DNA binding [27, 33]. I289 makes a van der Waals interaction with a dA base, and R290 makes hydrogen bond contacts with a dT base and a dG base (paired with a dC base in the consensus sequence of PBX1) (Fig. S3). The main difference between the DNA sequences recognized by TGIF1 (5′-TGTCA-3′) and PBX1 (5′-AATCAT-3′) is the nucleotides before the underlined dinucleotide TC. It is reasonable to consider that TGIF1 R221 plays a similar role to PBX1 R290, which recognizes the dinucleotide TC, while TGIF1 R220 may account for recognizing the different nucleotide before the dinucleotide TC, because instead of I289-mediated binding of the dA base in the PBX1-HD:DNA complex, R220 could serve as a hydrogen bond donor for the dG before the dinucleotide TC in the TGIF1-recognized DNA sequence. In addition, based on the model, it also can be seen that two HPE-related residues, P192 and R219, unlike R220 and R221, are not buried in the major groove without direct contact with DNA bases, but the possibility of their interaction with sugar-phosphate backbone could not be excluded. The model built here could provide structural insights for understanding the interaction mechanisms of TGIF1-HD in complex with consensus DNA. Further experimental complex structure may be obtained once buffer conditions are optimized for more stable DNA-bound TGIF1-HD.

4. Conclusions

This study characterized the specific DNA binding interface of human TGIF1-HD and its conformational exchanges upon DNA binding, overcoming difficulties in NMR titrations caused by NMR resonance disappearance and weak signal intensity. The middle of the α3 helix was suggested to be the specific DNA binding interface, and other regions, including the α1 and α2 helices, experienced conformational changes to some extent once the DNA was bound. Low populated DNA-bound complex in the slow conformational exchange regime under the conditions of this study might be associated with the TGIF1-HD function. Further functional assays of designed mutant proteins evidenced that residues R220 and R221 in the α3 helix are critical for DNA binding. This study provides new insight into the mechanism of DNA binding of TGIF1-HD in the aspects of the interaction interface, conformational dynamics, and key residues for DNA binding.

Conflicts of interest

There are no conflicts to declare.

Acknowledgments

This work was supported by funds from the National Key R&D Program of China (grant number 2016YFA051201), the National Natural Sciences Foundation of China (grant numbers 21575155 and 21735007), and the Program and Key Research Program of Frontier Sciences by the Chinese Academy of Sciences (grant number QYZDJ-SSW-SLH027). We thank Jing Fang (Waters Corp., Milford, MA, USA), Xi Chen (Waters, Shanghai) and Lan Kun Song (Waters, Shanghai) for providing HDX datasets and helping with data processing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbapap.2018.07.005.

References

[1] E. Bertolino, B. Reimund, D. Wildt-Perinic, R.G. Clerc, A novel homeobox protein which recognizes a TGT core and functionally interferes with a retinoid-responsive motif, J. Biol. Chem. 270 (1995) 31178–31188.
[2] R. Hamid, S.J. Brandt, Transforming growth-interacting factor (TGIF) regulates proliferation and differentiation of human myeloid leukemia cells, Mol. Oncol. 3 (2009) 451–463.
[3] K.W. Gripp, D. Wotton, M.C. Edwards, E. Roessler, L. Ades, P. Meinecke, A. Richieri-Costa, E.H. Zackai, J. Massague, M. Muenke, S.J. Elledge, Mutations in TGIF cause holoprosencephaly and link NODAL signalling to human neural axis determination, Nat. Genet. 25 (2000) 205–208.
[4] K.B. El-Jaick, S.E. Powers, L. Bartholin, K.R. Myers, J. Hahn, I.M. Orioli, M. Ouspenskaia, F. Lacbawan, E. Roessler, D. Wotton, M. Muenke, Functional analysis of mutations in TGIF associated with holoprosencephaly, Mol. Genet. Metab. 90 (2007) 97–111.
[5] C.P. Chen, S.R. Chern, S.H. Du, W. Wang, Molecular diagnosis of a novel heterozygous 268 C→T (R90C) mutation in TGIF gene in a fetus with holoprosencephaly and premaxillary agenesis, Prenat. Diagn. 22 (2002) 5–7.
[6] C. Aguilella, C. Dubourg, J. Attia-Sobol, J. Vigneron, M. Blayau, L. Pasquier, L. Lazaro, S. Odent, V. David, Molecular screening of the TGIF gene in holoprosencephaly: identification of two novel mutations, Hum. Genet. 112 (2003) 131–134.
[7] J. Zhu, S. Li, T.A. Ramelot, M.A. Kennedy, M. Liu, Y. Yang, Structural insights into the impact of two holoprosencephaly-related mutations on human TGIF1 homeodomain, Biochem. Biophys. Res. Commun. 496 (2018) 575–581.
[8] Z.M. Liu, J.T. Tseng, D.Y. Hong, H.S. Huang, Suppression of TG-interacting factor sensitizes arsenic trioxide-induced apoptosis in human hepatocellular carcinoma cells, Biochem. J. 438 (2011) 349–358.
[9] B.W. Yeh, W.J. Wu, W.M. Li, C.C. Li, C.N. Huang, W.Y. Kang, Z.M. Liu, H.S. Huang, Overexpression of TG-interacting factor is associated with worse prognosis in upper urinary tract urothelial carcinoma, Am. J. Pathol. 181 (2012) 1044–1055.
[10] M.Z. Zhang, O. Ferrigno, Z. Wang, M. Ohnishi, C. Prunier, L. Levy, M. Razzaque, W.C. Horne, D. Romero, G. Tzivion, F. Colland, R. Baron, A. Atfi, TGIF governs a feed-forward network that empowers Wnt signaling to drive mammary tumorigenesis, Cancer Cell 27 (2015) 547–560.
[11] D. Wotton, R.S. Lo, S. Lee, J. Massague, A Smad transcriptional corepressor, Cell 97 (1999) 29–39.
[12] D. Wotton, R.S. Lo, L.A. Swaby, J. Massague, Multiple modes of repression by the Smad transcriptional corepressor TGIF, J. Biol. Chem. 274 (1999) 37105–37110.
[13] T.A. Melhuish, D. Wotton, The interaction of the carboxyl terminus-binding protein with the Smad corepressor TGIF is disrupted by a holoprosencephaly mutation in TGIF, J. Biol. Chem. 275 (2000) 39762–39766.
[14] M. Sharma, Z. Sun, 5′TG3′ interacting factor interacts with Sin3A and represses AR-mediated transcription, Mol. Endocrinol. 15 (2001) 1918–1928.
[15] D. Wotton, P.S. Knoepfler, C.D. Laherty, R.N. Eisenman, J. Massague, The Smad transcriptional corepressor TGIF recruits mSin3, Cell Growth Differ. 12 (2001) 457–463.
[16] T.R. Burglin, M. Affolter, Homeodomain proteins: an update, Chromosoma 125 (2016) 497–521.
[17] T.R. Burglin, Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals, Nucleic Acids Res. 25 (1997) 4173–4180.
[18] M. Mojsin, J. Popovic, N. Kovacevic Grujicic, M. Stevanovic, TG-interacting factor (TGIF) downregulates SOX3 gene expression in the NT2/D1 cell line, J. Genet. Genomics 39 (2012) 19–27.
[19] C. Pramfalk, T.A. Melhuish, D. Wotton, Z.Y. Jiang, M. Eriksson, P. Parini, TG-interacting factor 1 acts as a transcriptional repressor of sterol O-acyltransferase 2, J. Lipid Res. 55 (2014) 709–717.
[20] A. Bahrami, A.H. Assadi, J.L. Markley, H.R. Eghbalnia, Probabilistic interaction network of evidence algorithm and its application to complete labeling of peak lists from protein NMR spectroscopy, PLoS Comput. Biol. 5 (2009) e1000307.
[21] P. Vallurupalli, G. Bouvignies, L.E. Kay, Studying "invisible" excited protein states in slow exchange with a major state conformation, J. Am. Chem. Soc. 134 (2012) 8148–8161.
[22] M. Guenneugues, P. Berthault, H. Desvaux, A method for determining B1 field inhomogeneity. Are the biases assumed in heteronuclear relaxation experiments usually underestimated? J. Magn. Reson. 136 (1999) 118–126.
[23] K.H. Frandsen, K.K. Rasmussen, M.R. Jensen, K. Hammer, M. Pedersen, J.C. Poulsen, L. Arleth, L. Lo Leggio, Binding of the N-terminal domain of the lactococcal bacteriophage TP901-1 CI repressor to its target DNA: a crystallography, small angle scattering, and nuclear magnetic resonance study, Biochemistry 52 (2013) 6892–6904.
[24] A. Proudfoot, H.L. Axelrod, M. Geralt, R.J. Fletterick, F. Yumoto, A.M. Deacon, M.A. Elsliger, I.A. Wilson, K. Wuthrich, P. Serrano, Dlx5 Homeodomain:DNA complex: structure, binding and effect of mutations related to split hand and foot malformation syndrome, J. Mol. Biol. 428 (2016) 1130–1141.
[25] C.S. Kenchappa, P.O. Heidarsson, B.B. Kragelund, R.A. Garrett, F.M. Poulsen, Solution properties of the archaeal CRISPR DNA repeat-binding homeodomain protein Cbp2, Nucleic Acids Res. 41 (2013) 3424–3435.
[26] A.N. Hoofnagle, K.A. Resing, N.G. Ahn, Protein analysis by hydrogen exchange mass spectrometry, Annu. Rev. Biophys. Biomol. Struct. 32 (2003) 1–25.
[27] D.E. Piper, A.H. Batchelor, C.P. Chang, M.L. Cleary, C. Wolberger, Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation, Cell 96 (1999) 587–597.
[28] A. Jolma, Y.M. Yin, K.R. Nitta, K. Dave, A. Popov, M. Taipale, M. Enge, T. Kivioja, E. Morgunova, J. Taipale, DNA-dependent formation of transcription factor pairs alters their binding specificity, Nature 527 (2015) 384–388.
[29] J.N. Foot, M. Feracci, C. Dominguez, Screening protein—single stranded RNA complexes by NMR spectroscopy for structure determination, Methods 65 (2014)

288–301.
[30] S.C. Chiliveri, M.V. Deshmukh, Recent excitements in protein NMR: large proteins and biologically relevant dynamics, J. Biosci. 41 (2016) 787–803.
[31] D.F. Hansen, P. Vallurupalli, P. Lundstrom, P. Neudecker, L.E. Kay, Probing chemical shifts of invisible states of proteins with relaxation dispersion NMR spectroscopy: how well can we do? J. Am. Chem. Soc. 130 (2008) 2667–2675.
[32] D. Long, C.B. Marshall, G. Bouvignies, M.T. Mazhab-Jafari, M.J. Smith, M. Ikura, L.E. Kay, A comparative CEST NMR study of slow conformational dynamics of small GTPases complexed with GTP and GTP analogues, Angew. Chem. Int. Ed. Engl. 52 (2013) 10771–10774.
[33] T. Sprules, N. Green, M. Featherstone, K. Gehring, Lock and key binding of the HOX YPWM peptide to the PBX homeodomain, J. Biol. Chem. 278 (2003) 1053–1058.