MATERIALS

Selection of buffer containing zinc ion

Although zinc phosphate $Zn_3(PO_4)_2$ is considered to be insoluble to water, we used 20 mM potassium phosphate buffer (pH6.0) containing 20 μM $ZnCl_2$, as described in the text. According to the solubility product of $Zn_3(PO_4)_2$, 9.0×10^{-33} M^5, and the pK values of phosphate, 7.20 and 12.38 (e.g., Dean, 1979), Zn^{2+} ion should be soluble up to 33 μM in 20 mM phosphate buffer at pH 6.0. In order to make it clearer that the phosphate buffer contains soluble zinc ion enough for the folding of the WRKY domain, we have compared the DNA-binding ability in different buffers by SPR (Fig. S4). The binding affinity in the phosphate buffer is similar or slightly more than that in a citrate buffer (pH6.0). Furthermore, in the phosphate buffer containing 1 mM EDTA instead of 20 μM $ZnCl_2$, the binding became much weaker. Therefore, zinc dependence of the DNA-binding activity was clearly observed in the phosphate buffer used in this study. Indeed, we have observed the zinc-dependent folding of the protein by CD and NMR using the same phosphate buffer system in the previous study (Yamasaki et al., 2005).

It should be noted that the concentration of the Zn^{2+} ion is much less than the protein concentration in the NMR buffer. It is very likely that the bound zinc is originated from the media for the cell-free protein synthesis, where 1 mM ZnSO4 was added (Matsuda et al., 2007). The affinity of Zn^{2+} to the protein should be tight enough not to be released in the buffers at lower zinc concentration.

In addition, dithiothreitol was added in the buffer used for NMR in order to avoid oxidation of the zinc-coordinating Cys residues, on a presumable occasion of the ion

release during the long measurements, but not in the buffer for SPR.

Structure determination

The structure of the complex was calculated for region Leu407–Ala469 of the protein and region [5'-CCTTTGACCA-3' / 5'-TGGTCAAAGG-3'] of the DNA. The distance constraints derived from the NOESY spectra at a mixing time of 100 ms were classified into four categories, 1.5–2.8, 1.5–3.5, 2.0–4.5, and 2.5–6.0 Å, according to the relationship that NOE intensity is inversely proportional to the sixth power of distance. The average intensities of the intra-residue NOEs of Hδ-Hε pairs of Phe or Tyr residues were used for the calibration of the intensities. Stereospecific assignments previously made for the free protein (Yamasaki *et al*, 2005) were employed unless significant peak shift (> 0.01 ppm) was observed. When a pair of NOEs from chemically equivalent protons (e.g., Hβ2 and Hβ3) without stereospecific assignment were classified into the same distance category, two separate constraints were imposed as if they were stereospecifically assigned. For other NOEs due to stereospecifically unassigned protons, the constraints were imposed using the sum-averaged distances from all the equivalent protons (Nilges, 1993). To maintain hydrogen bonds that were detected by the HSQC spectrum of the sample in D_2O, the same acceptors as for the free protein (Yamasaki *et al*, 2005) were assigned to the donors, and the donor-acceptor and proton-acceptor distances were restrained to 2.5–3.5 and 1.5–2.5 Å, respectively. A force constant of 75–150 kcal mol^{-1} $Å^{-2}$ was used for the above distance constraints.

To maintain the zinc coordination by Cys434, Cys439, His463, and His465,

theoretical constraints were imposed on the Zn–Sγ and Zn–Nδ/ε (Nδ for His463 and Nε for His465) distances, the Sγ–Zn–Sγ, Sγ–Zn–Nδ/ε, and Cβ–Sγ–Zn angles, and the Zn–imidazole ring planarities, as described (Yamasaki *et al*, 2005). These constraints were included only in the final part of the calculation (see below), and, in the earlier parts, other temporary distance constraints between the above Sγ and Nδ/ε atoms for 3.0–4.7 Å were instead introduced, to maintain orientations of the side-chains.

RDCs were restrained by a harmonic potential of 1.0 kcal mol^{-1} Hz^{-2}. Alignment tensor parameters, D_a and R (D_r/D_a) were determined by a grid search method (Clore *et al*, 1998), using number of accepted structures as an index (Fig. S5).

By using the $^{1}H_{\alpha}$, $^{1}H_{N}$, $^{13}C_{\alpha}$, and ^{15}N chemical shifts, constraints regarding backbone ϕ and ψ angles were generated by TALOS+ (Shen *et al*, 2009). Only the restraints that were predicted unambiguously, as judged by the program, and caused no significant violations during the structure calculation were employed. In addition, those with relevant ^{1}H-^{15}N HSQC cross peaks that shifted very largely by the DNA titration experiment, i.e., Lys416, Tyr417, Gln419, Tyr430, and Tyr431 (Yamasaki *et al*, 2005), were excluded, considering that chemical shifts of these residues do not properly reflect the protein conformation. A force constant of 5–400 kcal mol^{-1} rad^{-2} was used for the above dihedral angle restraints, depending on the stage of the simulated annealing (see below).

For DNA backbone, α, β, and ζ torsion angles were weakly restrained to ranges –60°±30°, 180°±30°, and –90°±30°, respectively, assuming that the DNA conformation stays within boundary of the B-form (Tjandra *et al*, 2000). In addition, ε torsion angles

were restrained to stereochemically allowed range, 225°±75° (Phan *et al*, 2005). These theoretical constraints were carefully introduced after examining that the DNA structures calculated without them are in the typical B-form (data not shown). Force constants were the same as TALOS+ dihedral angle restraints. In addition, sugar puckering was fixed to C2'-endo, by restraining the C1'–C2'–C3'–C4', C5'–C4'–C3'–C2', and C1'–O4'–C4'–C5' dihedral angles to –34.9°, –86.4°, and –106.4°, respectively, as implemented in the CNS program (Brünger *et al*, 1998). To maintain DNA base pairs, hydrogen bonds in the typical Watson-Crick base-paring were restrained in a manner similar to that for the protein hydrogen bonds. Also, planar constraints were applied for all nonhydrogen atoms included in the same base pair, by 150 kcal mol^{-1} $Å^{-2}$.

Structure calculation was carried out using CNS (Brünger *et al*, 1998) in essentially five consecutive parts. In the first part, a random simulated annealing (Nilges *et al*, 1998) was performed for the protein moiety without zinc ion, including NOEs, hydrogen bonds, the temporary distance constraints between zinc-coordinating atoms, and TALOS+ dihedral angle constraints, but not residual dipolar couplings. Torsion angle molecular dynamics of a high temperature annealing at a simulated temperature of 50,000 K for 1,000 steps and of a slow-cool annealing at initially 50,000 K for 1,000 steps were carried out, followed by a Cartesian molecular dynamics of a slow-cool annealing at initially 2,000 K for 3,000 steps. In the second part, the zinc-free protein structure and a B-DNA structure produced by the Insight II program (Accelrys) was combined by a rigid-body minimization protocol, including NOEs, hydrogen bonds, TALOS+ dihedral angle constraints, residual dipolar couplings, temporary distance

constraints between zinc-coordinating atoms, and theoretical constants regarding base pairs and backbone torsion angles of DNA. In the third part, simulated annealing for the zinc-free protein/DNA complex was performed by torsion angle dynamics, including the same constraints as in the second part. A constant temperature annealing at 5,000 K for 4,000 steps and a slow-cool annealing at initially 5,000 K for 4,000 steps were carried out. In the fourth part, simulated annealing for the zinc-bound protein/DNA complex was performed by Cartesian dynamics, including the same constraints as in the previous parts, except that the theoretical constraints regarding the zinc-coordination are included instead of the temporary distance constraints between the zinc-coordinating atoms. A constant temperature annealing at 2,000 K for 2,000 steps and a slow-cool annealing at initially 1,000 K for 2,000 steps were performed. In the final part, further simulated annealing was performed in the same way as in the fourth part, except that OPLS-AA (optimized potentials for liquid simulations-all atom) force field for electrostatic and van der Waals potentials (Jorgensen and Tirado-Rives, 1988) was applied with the same cut-off parameters as described by Fenn *et al*. (2011). This was followed by an energy minimization step without electrostatic forces and with the repel van der Waals functions that were applied in the previous parts.

From 100 initial structures, 20 structures with no distance violation larger than 0.2 Å, no torsion angle violation larger than 2 degrees, no residual dipolar coupling violation lager than 2 Hz, and the lowest total energies were selected as the final accepted structures. The minimized mean structures were produced by a protocol for the selection of the accepted structures in the CNS program.

The average RMSD of the co-ordinates from the unminimized mean structure of the ensemble (Table 1 in the main text) were calculated using the program MOLMOL (Koradi *et al*, 1996), after fitting the relevant atoms. The ϕ and ψ dihedral angles of protein were analyzed by using the program Procheck-NMR (Laskowski *et al*, 1996). The secondary structure elements were also identified by Procheck-NMR. The experimental constraints and stereochemical properties of the NMR solution structure of WRKY4-C are shown in Table 1.

References

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuzewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54,** 905–921

Clore, G. M., Gronenborn, A. M., and Tjandra, N. (1998) Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J Magn Reson* **131,** 159–162

Dean, J. A. (1979) Lange's HANDBOOK OF CHEMISTRY, 12th ed., McGRAW-HILL, New York, pp.5.12

Fenn, T. D., Schnieders, M. J., Mustyakimov, M., Wu, C., Langen, P., Pande, V. S., and Brünger, A. T. (2011) Reintroducing electrostatics into macromolecular crystallographic refinement: application to neutron crystallography and DNA

hydration. *Structure* **19**, 523–533

Jorgensen, W. L., and Tirado-Rives, J. (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc* **110**, 1657–1666

Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J Mol Graph* **14,** 51–55

Kraulis, P. J. (1991) *MOLSCRIPT*: a program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* **24,** 946–950

Laskowski, R. A., Rullman, J. A. C., MacArthur, W. M., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biol NMR* **8,** 477–486

Matsuda, T., Kigawa, T., Koshiba, S., Inoue, M., Aoki, M., Yamasaki, K., Seki, M., Shinozaki, K., and Yokoyama, S. (2007) Cell-free synthesis of zinc-binding proteins. *J Struct Funct Genomics* **7,** 93–100

Nilges, M. (1993) A calculation strategy for the structure determination of symmetric dimers by ^{1}H NMR. *Prot Struct Funct Genet* **17,** 297–309

Nilges, M., Clore, G. M., and Gronenborn, A. M. (1998) Determination of the three-dimensional structures of proteins from inter-proton distance data by dynamic simulated annealing from a random array of atoms. *FEBS Lett* **239,** 129–136

Phan, A. T., Kuryavyi, V., Ma, J. B., Faure, A., Andreola, M. L., and Patel, D. J. (2005) An interlocked dimeric parallel-stranded DNA quadruplex: A potent inhibitor of HIV-1 integrase. *Proc Natl Acad Sci USA* **102,** 634–639

Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS plus: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **22,** 213-223

Tjandra, N., Tate, S. I., Ono, A., Kainosho, M., and Bax, A. (2000) The NMR structure of a DNA dodecamer in an aqueous dilute liquid crystalline phase. *J Am Chem Soc* **122,** 6190–6200

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*. (New York: John Wiley & Sons, Inc.)

Yamasaki, K., Kigawa, T., Inoue, M., Tateno, M., Yamasaki, T., Yabuki, T., Aoki, M., Seki, E., Matsuda, T., Tomo, Y., Hayami, N., Terada, T., Shirouzu, M., Tanaka, A., Seki, M., Shinozaki, K., and Yokoyama, S. (2005) Solution structure of an Arabidopsis WRKY DNA binding domain. *Plant Cell* **17,** 944–956

LEGENDS FOR SUPPLEMENTARY FIGURES

Fig. S1: NOESY spectra of the complex of AtWRKY4-C and the 16-mer W-box DNA without modification (left panels) or the DNA with replacement of T9' by dU (right panels), recorded at 303 K, ^{1}H frequency of 750 MHz, and a mixing time of 100 ms. The protein and DNAs are unlabeled. Lines in red and cyan indicate resonances due to the protein and DNAs, respectively. NOEs among G8', T9', and G10' protons appearing in the left panels are typical of the connectivities between adjacent nucleotides (Wüthrich, 1996). Resonances of H8 of C8' and H6 of dU9' of the DNA with replacement of T9' by dU (right panels) are not identified presumably because the fluctuations amplified by the replacement caused local signal broadening. Note that resonances of the DNA protons in the molecular interface, e. g., H5 and H2' of C8', and H6, H7, and H2' of T9' in the left panels, are largely shifted (0.38, -0.60, -0.87, -1.35, and -0.46 ppm, respectively) from the average chemical shifts (5.60, 2.10, 7.36, 1.56, and 2.12 ppm, respectively, registered at BMRB; http://www.bmrb.wisc.edu/). These are conserved for H5 and H2' of C8' and H2' of dU9' in the DNA with the base replacement, suggesting that the binding geometry is not significantly altered by the replacement.

Fig. S2: The ensemble of the 20 selected structures of the complex of AtWRKY4-C and a 16-mer W-box DNA in stereo view. The structures are fitted by using the backbone nonhydrogen atoms of the protein (a) or all nonhydrogen atoms of the DNA (b). The presented region of the protein is Tyr412–His465, and that for DNA is C4-C11/G4'-G11'. The figures were produced by Molscript (Kraulis, 1991).

Fig. S3: A presumable hydrogen-bonding contact between Arg415 and G base at position 5 (stereo view). From the present complex structure, T5 was substituted by G and the side-chain torsion angles of Arg415 were altered. The relevant nucleotide and residue are shown by stick, while the other protein and DNA moieties are shown by wire in magenta and black, respectively. Protein backbone was indicated by ribbon in magenta. Insight II program (Accelrys) was used to produce this figure.

Fig. S4: Binding of AtWRKY4-C to a 16-mer W-box DNA observed by SPR. Equilibrium response values relative to the maximum response value expected when the protein molecules bind to all the immobilized DNA molecules at a 1:1 stoichiometry. The buffers used are 20 mM potassium phosphate buffer (pH 6.0) containing 100 mM KCl, 20 μM $ZnCl_2$, and 0.005% Tween20 (black, circle), 20 mM citrate buffer (pH 6.0) containing 100 mM KCl, 20 μM $ZnCl_2$, and 0.005% Tween20 (red, square), and 20 mM potassium phosphate buffer (pH 6.0) containing 100 mM KCl, 1 mM EDTA, and 0.005% Tween20 (blue, square). Fitting curves to the simple 1:1 binding model are indicated with binding constants shown alongside the lines.

Fig. S5: A grid search for determination of alignment tensor: number of accepted structures from 50 trials, without violations more than 0.2 Å, 2°, 2 Hz of constraints for NOEs, dihedral angles, and residual dipolar couplings, respectively, as functions of parameters for alignment tensor, D_a (axial components) and R (rate of rhombic

component, D_r, to D_a). Each trial calculation was performed as described in the above section for "structure determination", except that the last part using the OPLS-AA force field was not included. In addition, 1.2 % of the total NOEs, which are mostly intermolecular ones, were assigned after determination of the tensor parameter. Numbers of accepted structures without limitation of the overall energy (a) or with limitation of 200 kcal mol^{-1} (b) are shown. Red, orange, and yellow colors show that the accepted structures are ≥15, ≥10, and ≥5, respectively. According to the present profile, D_a of −15.0 Hz and R of 0.4 were selected for the final structure determination.

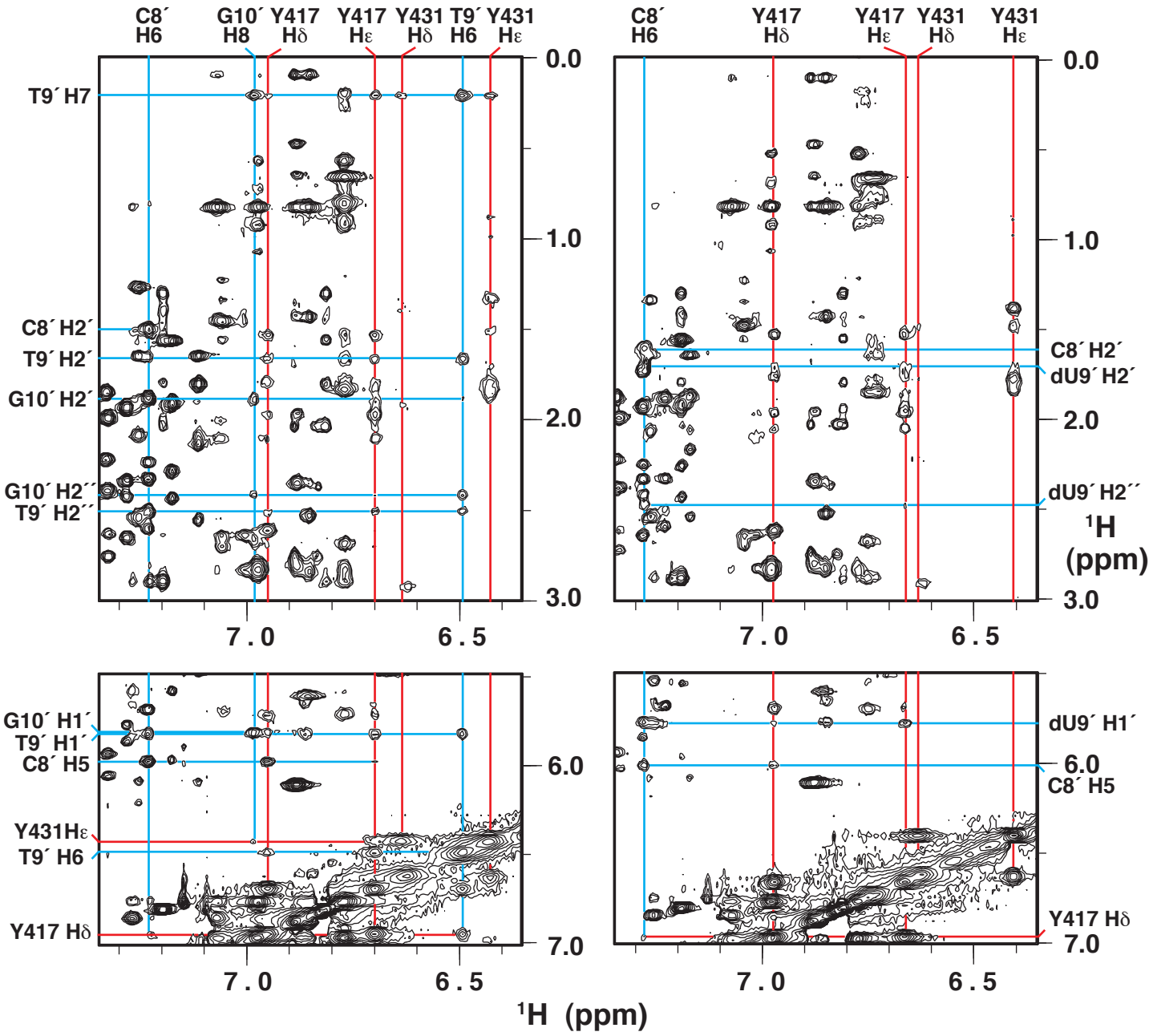
C8´ H6
G10´ H8
Y417 Hδ
Y417 Hε
Y431 Hδ
T9´ H6
Y431 Hε
T9´ H7
C8´ H2´
T9´ H2´
G10´ H2´
G10´ H2´´
T9´ H2´´
0.0
1.0
2.0
3.0
7.0
6.5
C8´ H6
Y417 Hδ
Y417 Hε
Y431 Hδ
Y431 Hε
C8´ H2´
dU9´ H2´
dU9´ H2´´
¹H (ppm)
G10´ H1´
T9´ H1´
C8´ H5
Y431Hε
T9´ H6
Y417 Hδ
6.0
7.0
dU9´ H1´
C8´ H5
¹H (ppm)

Fig. S1

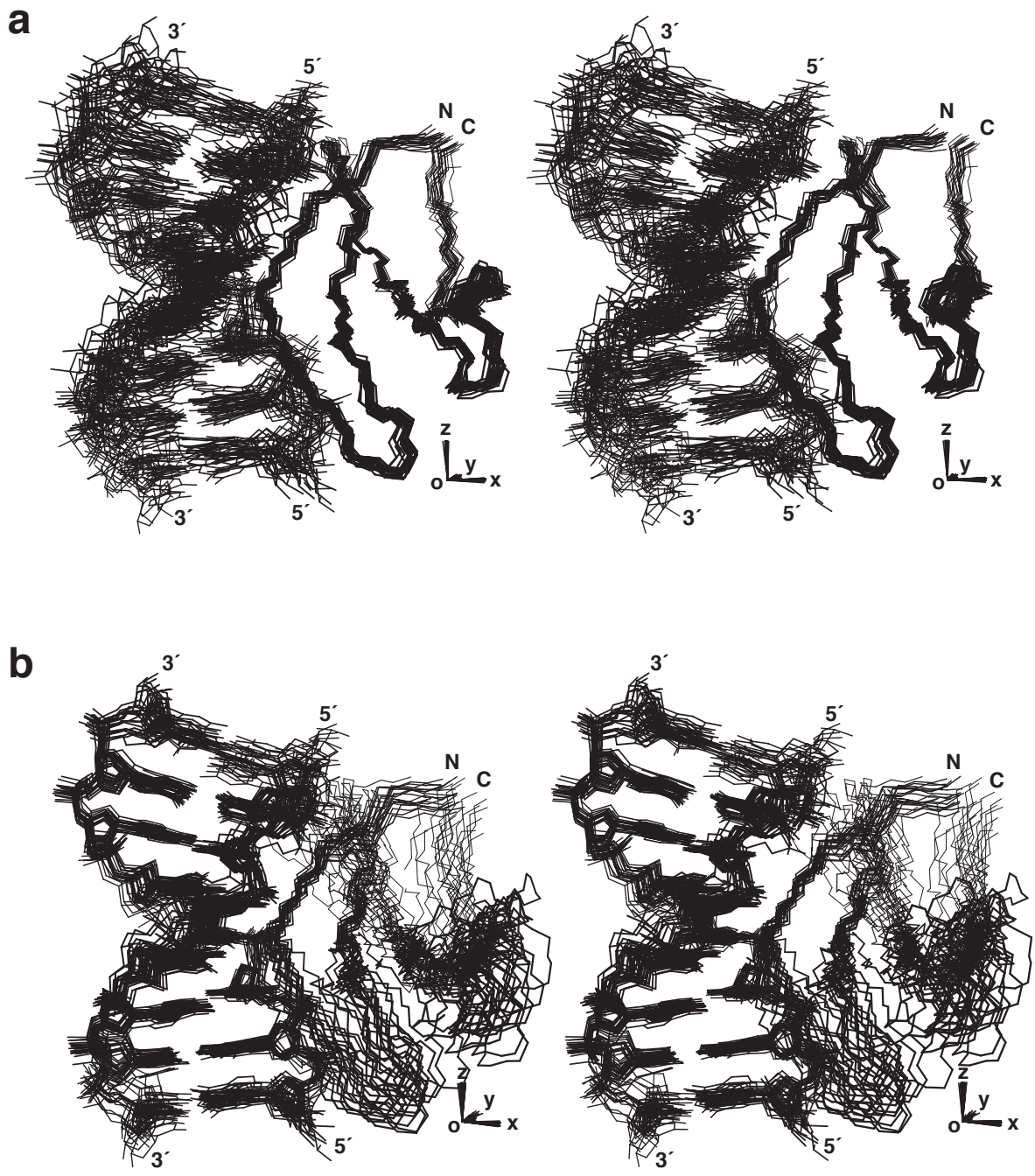

Fig. S2

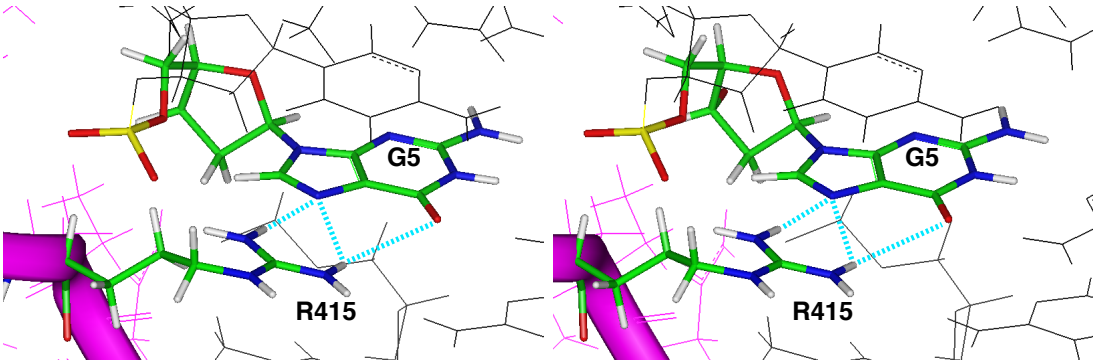

Fig. S3

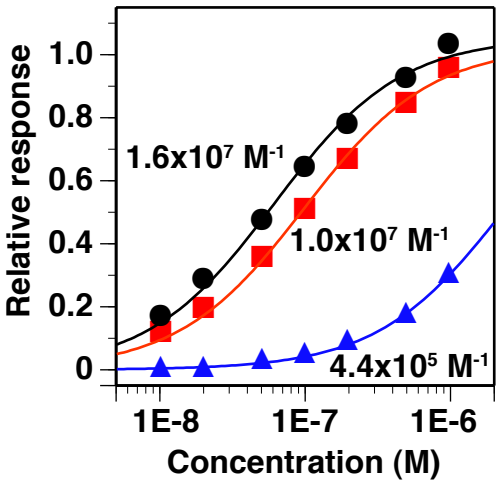
Relative response
1.0
0.8
0.6
0.4
0.2
0
1.6x10⁷ M⁻¹
1.0x10⁷ M⁻¹
4.4x10⁵ M⁻¹
1E-8
1E-7
1E-6
Concentration (M)

Fig. S4

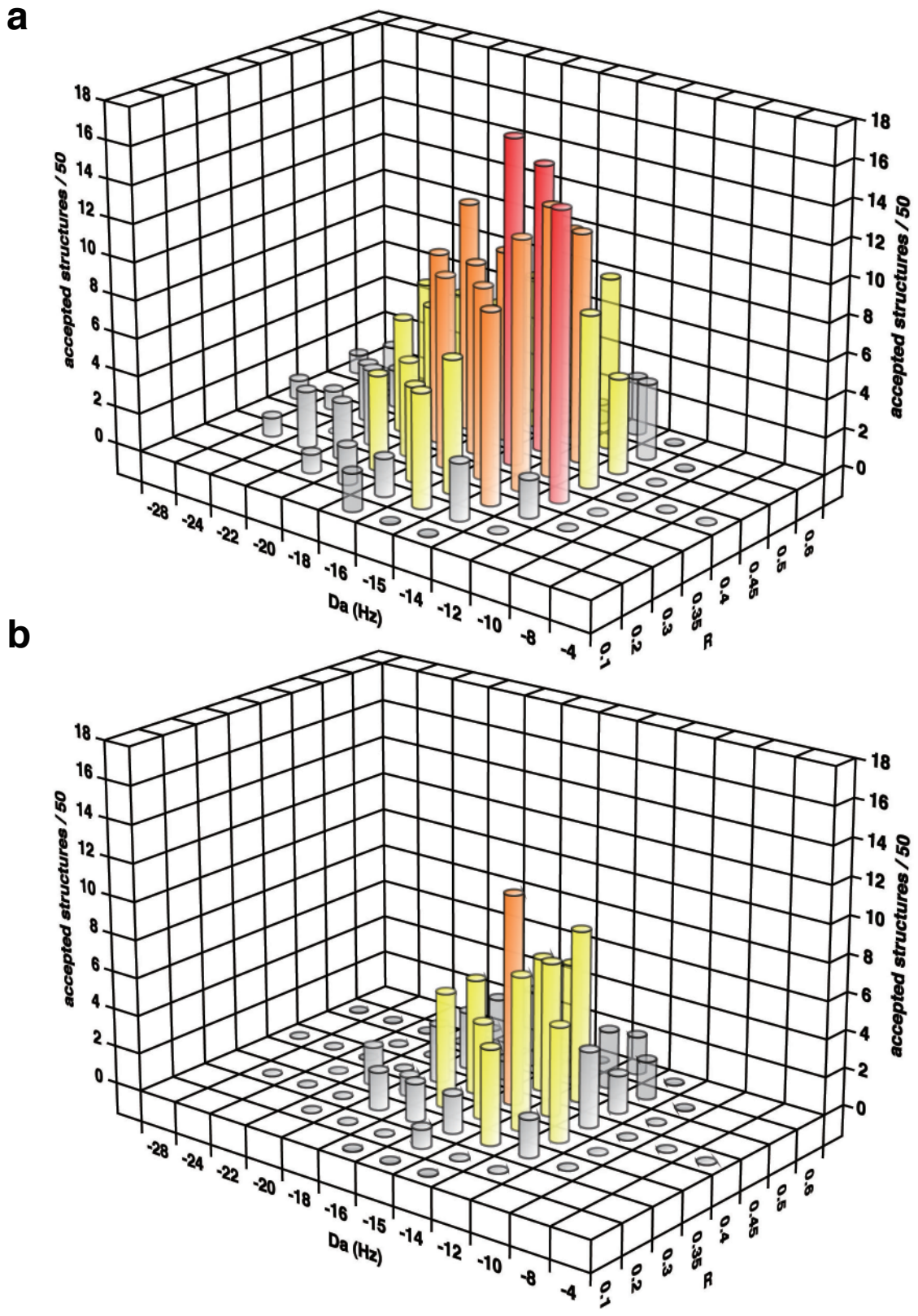
a
b
accepted structures / 50
Da (Hz)
R
0
2
4
6
8
10
12
14
16
18
-28
-24
-22
-20
-18
-16
-15
-14
-12
-10
-8
-4
0.1
0.2
0.3
0.35
0.4
0.45
0.5
0.6

Fig. S5