GATA1 Binding Kinetics on Conformation-Specific Binding Sites Elicit Differential Transcriptional Regulation

Atsushi Hasegawa,^{a,b} Hiroshi Kaneko,^{a,c} Daishi Ishihara,^{a,b} Masahiro Nakamura,^d Akira Watanabe,^d Masayuki Yamamoto,^{b,c} Cecelia D. Trainor,^e Ritsuko Shimizu^{a,c}

Department of Molecular Hematology, Tohoku University Graduate School of Medicine, Sendai, Japan^a; Department of Medical Biochemistry, Tohoku University Graduate School of Medicine, Sendai, Japan^b; Tohoku Medical Mega-Bank Organization, Tohoku University, Sendai, Japan^c; Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan^d; Laboratory of Molecular Biology, NIDDK, National Institutes of Health, Bethesda, Maryland, USA^e

GATA1 organizes erythroid and megakaryocytic differentiation by orchestrating the expression of multiple genes that show diversified expression profiles. Here, we demonstrate that GATA1 monovalently binds to a single GATA motif (Single-GATA) while a monomeric GATA1 and a homodimeric GATA1 bivalently bind to two GATA motifs in palindromic (Pal-GATA) and direct-repeat (Tandem-GATA) arrangements, respectively, and form higher stoichiometric complexes on respective elements. The amino-terminal zinc (N) finger of GATA1 critically contributes to high occupancy of GATA1 on Pal-GATA. GATA1 lacking the N finger-DNA association fails to trigger a rate of target gene expression comparable to that seen with the wild-type GATA1, especially when expressed at low level. This study revealed that Pal-GATA and Tandem-GATA generate transcriptional responses from GATA1 target genes distinct from the response of Single-GATA. Our results support the notion that the distinct alignments in binding motifs are part of a critical regulatory strategy that diversifies and modulates transcriptional regulation by GATA1.

One of the fundamental questions in the study of vertebrate gene expression regulation is how one transcription factor coordinately regulates the spatiotemporal expression of multiple target genes. The function of transcription factors appears to be modified by multiple mechanisms, including posttranslational modifications, alternative splicing isoforms, higher-order complex formation with diversified cofactors, and interaction with other transcription factors that recognize specific sequences located in adjacent and/or distant loci of the genome. Accumulating lines of evidence indicate that multiple mechanisms elaborately work in a complex manner to modify transcription factor activity, which enables the transcription factor to exert diversified gene expression regulation. We have been characterizing GATA1 that coordinately regulates multiple target genes during the development and differentiation of erythroid and megakaryocytic lineages through binding to GATA motif (A/T)GATA(A/G) (1, 2).

GATA1 is the founding member of the GATA family of transcription factors, which is composed of 6 members in vertebrates. GATA1 and GATA2 are partially overlapping in their expression profiles during erythroid and megakaryocyte differentiation and regulate each other's expression. In hematopoietic stem/progenitor cells, *Gata1* gene expression is repressed by an epigenetic mechanism that precludes access by GATA2 (3). However, once GATA2 activates *Gata1* gene expression, GATA1 self-activates its own gene and conversely represses *Gata2* gene expression. Consequently, the expression levels of GATA1 and GATA2 show a dynamic feature, referred to as "GATA-factor switching" during erythroid development (4). GATA1 and GATA2 share binding to their target genes and redundantly and/or competitively regulate the genes (5). GATA1 receives posttranslational modifications, including phosphorylation, ubiquitination, sumoylation, and acetylation, to increase its functional diversity (6–9), although a comprehensive understanding of how these modifications are integrated into the regulatory activities of GATA1 and/or of the

existence of similar modifications to GATA2 awaits further studies.

GATA1 has four functional domains, i.e., two transactivation domains residing in amino and carboxyl termini, which transactivate GATA1 target genes redundantly and/or cooperatively (10), and two zinc finger domains in the middle of the protein. The carboxyl-terminal zinc (C) finger is essential for the DNA binding of GATA1, whereas the amino-terminal zinc (N) finger retains insufficient binding activity to GATA motifs by itself but contributes to stabilize the binding of the C finger to a double GATA site arranged in a palindromic manner (11, 12). Of note, while this two-finger structure is conserved in six distinct vertebrate GATA factors, there exist GATA factors with a single zinc finger in non-vertebrates, indicating that only the C finger and the following basic tail region are evolutionarily conserved in both vertebrate and nonvertebrate GATA factors (13). In our transgenic rescue analyses, GATA1 lacking the N finger supports embryonic erythropoiesis in mice up to embryonic day 15.5 (E15.5) but does not support definitive erythropoiesis. However, mice without the C finger die *in utero* by E12.5, showing a phenotype similar to that of the parental GATA1-deficient mice without rescue (14). There-

Received 9 January 2016 Returned for modification 3 February 2016
Accepted 17 May 2016

Accepted manuscript posted online 23 May 2016

Citation Hasegawa A, Kaneko H, Ishihara D, Nakamura M, Watanabe A, Yamamoto M, Trainor CD, Shimizu R. 2016. GATA1 binding kinetics on conformation-specific binding sites elicit differential transcriptional regulation. *Mol Cell Biol* 36:2151–2167. doi:10.1128/MCB.00017-16.

Address correspondence to Cecelia D. Trainor, ceceliat@mail.nih.gov, or Ritsuko Shimizu, rshimizu@med.tohoku.ac.jp.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/MCB.00017-16>.

Copyright © 2016, American Society for Microbiology. All Rights Reserved.

fore, the zinc fingers of GATA1 have distinct functions and the role of the N finger is assumed to have been acquired during the molecular evolution of vertebrates.

A number of coactivators and corepressors have been reported to associate with GATA1. An intriguing feature of the cofactor binding to GATA1 is that the cofactor associations appear to be irrelevant for transcriptional activation or repression. For instance, FOG1 (friend of GATA1) shows a specific association with the N finger of GATA1 (15). While FOG1 recruits the nucleosome remodeling and deacetylase (NuRD) complex (16), FOG1 occupies most of the GATA1 binding sites irrespective of whether the genes are activated or repressed by GATA1 (17, 18). Similarly, GATA1 interacts with CREB-binding protein (CBP), but CBP is colocalized with GATA1 on the genes both activated and repressed by GATA1 (6, 19, 20). MED1, a component of mediator complex, acts as a coactivator of GATA1, but the mediator complex is found at the promoter regions of genes both activated and repressed by GATA1 (21).

It has been found that binding motifs for other transcription factors are differentially enriched adjacent to GATA motifs at both the GATA1-activated and -repressed genes (22) and this differential enrichment occurs at erythroid-specific and megakaryocyte-specific genes (23). Therefore, site-specific recruitment of the corresponding transcription factors likely alters the gene expression mediated by GATA1. Furthermore, if the given transcription factor is able to assemble into a complex with GATA1, allosteric modulation of GATA1 may result, as seen with the GATA1-LMO2-LDB1-SCL/TAL1 complex binding to E-box-GATA motifs (24, 25). This suggests that the sequence arrangement surrounding the GATA1 motif determines the gene expression outcome.

If we turn our attention to *cis*-acting GATA motifs, various erythroid genes have been shown to harbor functional GATA motifs. One of the best-characterized erythroid genes analyzed is the *Gata1* gene, which is equipped with two *cis*-acting regulatory domains containing GATA motifs, i.e., a single GATA motif in the GATA1 hematopoietic enhancer and a double GATA motif in the upstream promoter region (26). In the course of these *cis*-element analyses by means of reporter transgenic assays, we have noticed diversity in the transcriptional response to GATA1 expression that varies depending on the *cis*-element configuration. In fact, these *cis*-acting GATA elements are conserved in the human *GATA1* gene (26), and a specific mode of GATA1 binding to a palindromic GATA box has been noticed (11, 12).

In this study, we have examined GATA motif configuration-specific modulation of GATA1 function by using composite GATA elements in which two GATA motifs are aligned side by side, in either tandem or palindromic orientation. We have defined changes in the GATA1 binding and transactivation activity in accordance with the arrangement of *cis*-acting GATA motifs. We have demonstrated for the first time that GATA1 bivalently binds to the palindromic and tandem GATA motifs by using both the N and C fingers in the GATA1 monomer and two C fingers in GATA1 homodimer, respectively, while GATA1 binds monovalently to the single GATA motif. These diversities in binding of GATA1 to single, tandem, and palindromic GATA motifs lead to modulation in the transactivation activity of GATA1 and, in conjunction with our chromatin immunoprecipitation (ChIP) data, indicate the potential existence of many distinctly structured *cis*-acting GATA elements that differentially influence GATA1 activity.

MATERIALS AND METHODS

Generation of stable transfectant cell clones. HEK293T and mouse erythroleukemia (MEL) cells were cultured in high-glucose Dulbecco's modified Eagle's medium (Wako) and Roswell Park Memorial Institute medium (Wako), respectively, supplemented with 10% fetal bovine serum (Gibco) and 100 IU/ml of penicillin-streptomycin (Gibco) in 5% CO₂ at 37°C. Transfection was performed with either Lipofectamine LTX (Invitrogen Life Technologies) or Neon transfection system (100-μl tip) (Invitrogen Life Technologies) according to the manufacturer's protocols. For generation of HEK293T reporter clones, Single-GATA- and Pal-GATA-driven reporter constructs were generated by inserting fragments of 5'-CTCCGGCAACAGATAAGGAATCCCTG and 5'-TCCGGCAACC TATCAGATAAGGAATCCCTG standing in triplicate into the multiple cloning site of the pNL1.3 vector (Promega) just upstream of a minimal-promoter sequence derived from the pGL4.28 vector (Promega), followed by introduction into HEK293T cells. For generation of MEL clones that expressing biotinylated GATA1 proteins, a parental MEL clone stably carrying BirA-expressing plasmid (kindly provided by Alan Cantor, Children's Hospital and Dana Farber Cancer Institute, Harvard Medical School) was established, and subsequently pEF1α-FLAG-Bio-*G1*^{WT} and -*G1*^{R216Q} expression vectors were introduced as previously described (27). Transduced HEK293T and MEL cells were purified with 200 μg/ml hygromycin B (Invitrogen Life Technologies) and with 10 μg/ml puromycin (Gibco) plus 400 μg/ml G418 (Gibco), respectively, followed by isolation of single-cell clones by limiting dilution.

Protein quantifications. MEL cells and fetal liver cells were subjected to lysis in Laemmli lysis buffer with 2-mercaptoethanol, followed by SDS-PAGE in 15% and 10% polyacrylamide gels, respectively. After electrophoresis, the separated proteins were transferred to an Immobilon-P membrane (Millipore). Endogenous proteins were probed with primary anti-GATA1 or anti-lamin B antibodies (Santa Cruz Biotechnology), followed by detection of the primary globulins with horseradish peroxidase (HRP)-conjugated secondary antibodies (Invitrogen-Life Technologies). Biotinized proteins were detected with the Elite ABC standard kit (Vectastain) according to the manufacturer's protocol. Signals were visualized using ECL-Prime Western blotting detection reagents (GE Healthcare). Binding signal values were quantified using ImageJ software.

Recombinant proteins. Maltose binding protein (MBP)-fused GATA1 and GATA1 mutants, and glutathione S-transferase (GST)-fused GATA1 and GATA1 mutants were synthesized as previously described (10). The recombinant proteins were purified by affinity chromatography utilizing a Profinia instrument (Bio-Rad) according to the manufacturer's protocol with slight modifications.

SPR analyses. Surface plasmon resonance (SPR) analyses were performed with a Biacore-X100 (GE Healthcare). Double-stranded DNA probes in which the 5' ends of sense strands were biotinylated were purified by native PAGE. Subsequently, each probe was immobilized in one of two streptavidin-coated flow cells in a Sensor chip-SA (GE Healthcare) as an active flow cell of 1,000 resonance units (RU) DNA according to the manufacturer's protocol. The other flow cell was left blank for reference subtraction. Details of the running and regenerating conditions were as described previously (28). Data processing was performed with Biacore-X100 evaluation software (GE Healthcare). Sequences of DNA probes are described in Fig. 1A.

EMSAs. Electrophoresis mobility shift assays (EMSAs) were performed as previously described (10). Briefly, recombinant MBP fusion proteins were incubated with radioisotope-labeled oligonucleotides in the binding buffer containing 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 4% Ficoll, 1 mM dithiothreitol (DTT), and 75 mM KCl until the reaction mixture reached an equilibrium. A 100-fold excess of unlabeled competitor DNA was added into EMSA reaction tubes, and aliquots were applied at the indicated time points (minutes) to a progressively running gel. The gels were vacuum dried before the phosphorimager analysis. Binding signal values were quantified using ImageJ software and plotted as the percentage of complex shifted over the initial complex formed before com-

FIG 1 Analysis of GATA1 binding to Single-GATA. (A) Alignment of the double-stranded DNA probes fixed on the sensor chip. GATA motifs are depicted in pink. Arrows indicate the 5' to 3' orientation of consensus GATA motifs. (B) Schematic diagram of the domain structure of GATA1 and GATA1 mutants. N-TAD, amino-terminal transactivating domain; C-TAD, carboxyl-terminal transactivating domain; NF, N finger; CF, C finger. Arginine (R) 216 residue is replaced by glutamine (Q) and tryptophan (W) in G1^{R216Q} and G1^{R216W}, respectively, whereas three lysine residues (K45, K246, and K312) are substituted for alanine in G1^{3KA}. G1^{ΔNF} and G1^{ΔCF} lack N finger and C finger, respectively. (C, D) SPR analysis of GATA1 binding to Single-GATA. The influences of flow rate in the indicated protein concentrations (C) and kinetics parameters (D) of the respective GATA1 proteins on the binding to Single-GATA are shown. Measured data and regression analysis-fits are shown in red and black curves, respectively, in panel C. Typical results acquired in two independent trials are shown. (E) Off-rate analyses of GATA1 and R216-substituted mutants bound to Single-GATA. Free RI-labeled probe was loaded into lanes labeled P. Retarded complex containing GATA1 proteins and free probe are indicated by filled and open arrowheads, respectively. Scatchard plots are displayed on the right. Approximation lines in Scatchard plots of G1^{WT}, G1^{R216Q}, and G1^{R216W} were depicted in black, solid gray, and dotted gray lines, respectively. k_d values were derived from the slopes. Data are representative of two independent experiments. (F) Schematic diagram of predicted binding models of GATA1 to Single-GATA.

petition versus the time of competition. The sequences of the oligonucleotide probes are described in Fig. 1A.

Sequential oligonucleotide pulldown assays. Recombinant MBP fusion $G1^{WT}$, $G1^{R216Q}$, $G1^{R216W}$, $G1^{3KA}$, $G1^{DNF}$, and $G1^{ACF}$ and each of the corresponding GST fusion GATA1 proteins were incubated together with 5'-biotinylated double-stranded oligonucleotide probes in EMSA buffer for 30 min at room temperature. Protein-DNA complexes containing MBP fusion GATA1 protein were purified by amylose-resin (New England BioLabs) (the first pulldown). The purified complexes were then mixed with Pierce glutathione magnetic beads (Thermo Scientific), and subsequently protein-DNA complexes containing GST fusion GATA1 were precipitated (the second pulldown). Elutions obtained in the first and the second pulldown steps were loaded on native PAGE gels and transferred to Zeta-Probe GT genomic tested blotting membranes (Bio-Rad). After UV-cross-linking and blocking with 1% bovine serum albumin (BSA), biotinylated probes were detected by HRP-conjugated streptavidin (Dako) and the ECL-plus Western blotting detection system (GE Healthcare). The sequences of the oligonucleotide probes are described in Fig. 1A.

ChIP and ChPD-seq assays. Sample preparations were performed as previously described with minor modifications (29). Briefly, 1×10^7 liver cells of E12.5 mouse embryos and 1×10^7 cells of MEL cell clones that express biotinylated GATA1 proteins were cross-linked with 1% and 0.4% formaldehyde for 10 min, respectively, and neutralized with glycine. After sonication, endogenous and biotinylated GATA1-bound DNA fragments were precipitated with anti-GATA1 N6 antibody (SantaCluz Biotechnology) and Dynabeads M-280 streptavidin (Invitrogen Life Technologies), and subsequently precipitated DNA fragments were processed for quantitative PCR analysis with primers described in Table S4 in the supplemental material and high-throughput sequencing with the HiSeq system (Illumina), respectively. Chromatin pulldown sequencing (ChPD-seq) data were processed with BWA (Burrows-Wheeler alignment tool) and mapped into the MM9 mouse genome reference. GATA1-occupied regions were identified using MACS (Model-based Analysis of ChIP-seq) software. $TATCX_nGAT$ ($n = 1$ to 5) was defined as the Pal-GATA sequence. Single-GATA and Pal-GATA sequences were searched by FIMO software (30) with a parameter ('-thresh 0.05 -text -bgfile <background file>') to correct the distortion of background sequence frequencies to even (A:C:G:T = 0.25/0.25/0.25/0.25). ChPD-seq profiles in MEL cells were visualized on the GenomeJACK browser.

Luciferase assays. For the transient reporter/effector cotransfection/transactivation assays, reporter constructs carrying *Lmo2* and *Uros* gene enhancers were made by introducing a Pal-GATA-containing *Lmo2* gene enhancer (−11215/−10809) and *Uros* gene enhancer (−1157/−926) into the multiple cloning site of pGL4.10 vector (Promega) just upstream of a minimal-promoter sequence derived from pGL4.28 vector (Promega), respectively. HEK293T cells were washed twice with phosphate-buffered saline (PBS) and resuspended in 500 μ l of fresh medium at the concentration of 1.5×10^5 /ml, followed by transfection with 200 ng of reporter construct and 10 ng of *Renilla Luc* expression vector together with $G1^{WT}$ or $G1^{R216Q}$ expression vector at the indicated dosages. The total amount of DNA was adjusted to 300 ng with empty vector. After 24 h of transfection, luciferase activity was measured in triplicates at each dosage point of effector expression vectors with the Dual-Luciferase reporter assay system (Promega) with Lumat LB 9507 (Berthold Technologies) according to the manufacturer's protocol. The firefly luciferase value of each sample was normalized to its renilla luciferase value. For luciferase assays using stable HEK293T reporter cell lines, cells were washed twice with PBS and resuspended in 500 μ l of fresh medium at the concentration of 1.5×10^5 /ml, followed by transfection with $G1^{WT}$ or $G1^{R216Q}$ expression vector at the indicated dosages. The total amount of expression vector was adjusted to 200 ng with empty vector. Five microliters of cultured medium was taken at the indicated time points and used for measurement of luciferase activity with the Nano-Glo luciferase assay system (Promega) with the PHERAstar multiplate reader (BMG Labtech) according to the manufac-

turer's protocol. Luciferase value was defined as the subtraction of luciferase activity measured in the time point-matched cultured medium of untransfected cells from that measured in the respective medium of transfected cells.

Animals. Experimental procedures for animals were approved by the Institutional Animal Experiment Committee of the Tohoku University. Experiments were carried out in accordance with the Regulation for Animal Experiments in Tohoku University. $G1BAC-G1^{WT}$ and $G1BAC-G1^{R216Q}$ transgenic mouse lines were established as previously described with modifications (31). Briefly, the targeting vector for $G1BAC-Gata1^{WT}$ transgene was constructed by inserting a 5'-homologous region that contains a part of the first intron 2.0 kbp up to exon II (SacI site), a GATA1 gene, and a poly(A) signal, which were derived from the G1HRD-GATA1, into a vector containing a neomycin resistance (*Neo*) gene cassette, followed by insertion of a 1.3-kbp EcoRV-KpnI fragment that includes exon VI as a 3'-homologous region. For construction of target vector for $G1BAC-G1^{R216Q}$ transgene, a point mutation was introduced at R216 according to a PCR-based mutation protocol. RP23-443E19-based recombinant bacterial artificial chromosome (BAC) clones were established by homologous recombination as previously described (31) and were microinjected into fertilized eggs. For the analysis of transgene-derived *Gata1* mRNA levels, total RNA was isolated from bone marrow cells using Isogen or Isogen-LS (Nippon Gene). First-strand cDNA was synthesized with SuperScript III reverse transcriptase (Invitrogen Life Technologies) according to the manufacturer's protocol. The sequences of primers were 5'-GCTGAATCCTCTGCATCAAC (forward) and 5'-TAGGCCTCAGCTTCTCTGTA (reverse).

Peripheral blood analyses and colony assays. Mice were bled from the retro-orbital plexus. Hematopoietic indices were measured by means of a hemocytometer (Nihon Koden). The megakaryocyte CFU assay was performed using the MegaCult-C kit (Stem Cell Technologies) as previously described (32).

Microscopic and image analyses. E12.5 fetal livers were fixed, sectioned, and stained with hematoxylin and eosin solutions as previously described (29). Peripheral blood smears were stained with Wright-Giemsa solutions. Microscope images were captured with the BX51 microscope system and a DP73 charge-coupled-device (CCD) camera using the manufacturer's cellSens Standard software (all from Olympus). Images of fetus and fetal livers were taken with MZFL III stereomicroscopes (Leica) and a DP73 CCD camera. Pictures of mice were taken with μ TOUGH-8000 (Olympus).

Fluorescence-activated cell sorter (FACS) analyses. Mononuclear cells were stained with a combination of fluorescein isothiocyanate (FITC)-conjugated anti-CD71, phycoerythrin (PE)-conjugated anti-Ter119, and allophycocyanin (APC)-conjugated anti-c-kit antibodies (all from eBiosciences) or a combination of FITC-conjugated anti-CD41, PE-conjugated anti-CD61 (BD Bioscience Pharmingen), and APC-conjugated anti-c-kit antibodies for 30 min on ice. FITC-, PE-, and APC-conjugated rat IgG were used as isotype-matched controls. Dead cells were excluded by propidium iodide (2 μ g/ml). Cell populations were analyzed with BD FACSCalibur (Becton Dickinson).

Statistical analyses. Statistical analyses were performed with Student's *t* test. In Fig. 6B and 13B, data were analyzed employing the chi-square test and the Mann-Whitney rank sum test, respectively.

RESULTS

SPR analysis of GATA1 binding to Single-GATA. To elucidate how GATA1 binding is modified by the configuration of GATA motifs on DNA, we conducted SPR (surface plasmon resonance) analyses. We used a single GATA motif (Single-GATA), two GATA motifs aligned in palindromic and reversely palindromic orientations (Pal-GATA and rPal-GATA, respectively), and two GATA motifs aligned in series (Tandem-GATA), as depicted in Fig. 1A. Maltose binding protein (MBP)-fused recombinant proteins of wild-type GATA1 ($G1^{WT}$) and a series of GATA1 mutants

FIG 2 Analyses of GATA1 binding to Pal-GATA. (A, B) A comparison of the influences of flow rate in indicated protein concentrations (A) and kinetics parameters (B) of the respective GATA1 proteins on the binding to Pal-GATA. Typical results acquired in two independent SPR analyses are shown. For details, see the legend to Fig. 1. (C, D) Off-rate analyses of G1^{R216Q} (C) and G1^{R216W} (D) in comparison with GATA1 bound to Pal-GATA. Data are representative of two independent experiments. For details, see the legend to Fig. 1. (E) Schematic diagrams of predicted binding model of GATA1 to Pal-GATA.

(G1^{R216Q}, G1^{R216W}, G1^{ΔNF}, and G1^{3KA}) (Fig. 1B) were prepared as analytes of SPR. Of the GATA1 mutant proteins used, G1^{R216Q}, G1^{R216W}, and G1^{ΔNF} lacked the N finger-DNA interaction and G1^{R216Q} and G1^{R216W} maintained interaction with FOG1 (33, 34), while G1^{ΔNF} and G1^{3KA} lost the ability to generate GATA1 homodimers (33, 35–37). We immobilized each of the biotin-labeled oligonucleotides on a sensor chip and flushed these GATA1 mutant proteins of different concentrations across the sensor chip. Subsequently, we flushed protein-free buffer to retrieve the dissociated proteins from the sensor chip.

We found that the parameters of association rate (k_a), dissociation rate (k_d), equilibrium binding constants (K_D), and maximum binding capacity (R_{max}) were equivalent between G1^{WT} and GATA1 mutants when Single-GATA1 oligonucleotide was used as a ligand (Fig. 1C and D). Furthermore, fitting of the data was well correlated with the 1:1 binding model (Fig. 1D). Showing very

good agreement with the SPR data, off-rate analysis conducted by electrophoresis mobility shift assay (EMSA) showed that k_d values of G1^{R216Q} and G1^{R216W} for Single-GATA probe were similar to that of G1^{WT} (Fig. 1E). These results thus indicate that GATA1 binds to Single-GATA in a monovalent way via C finger without the influence of N finger (Fig. 1F).

GATA motif configuration diversifies GATA1-DNA binding modes. We then analyzed wild-type and mutant GATA1 binding to Pal-GATA. We found that the SPR parameters of G1^{WT} and Pal-GATA binding were different from those obtained from the G1^{WT} and Single-GATA binding, as they did not fit the 1:1 binding model (Fig. 2A and B). In contrast, G1^{R216Q}, G1^{R216W}, and G1^{ΔNF} mutants showed SPR binding parameters similar to those of the 1:1 binding model, as these mutants lack the N finger itself or the capability of N finger-DNA interaction. The N finger of the G1^{3KA} mutant binds DNA normally, so that the G1^{3KA}-Pal-GATA inter-

FIG 3 SPR analyses of GATA1 binding to rPal-GATA and Tandem-GATA. The influences of flow rate in the indicated protein concentrations (A, D) and kinetics parameters (B, E) of the respective GATA1 proteins on the binding to rPal-GATA (A, B) and Tandem-GATA (D, E) are shown. Data are representative of two independent experiments. For details, see the legend to Fig. 1. (C and F) Schematic diagrams of predicted binding models of GATA1 to rPal-GATA (C) and Tandem-GATA (F).

action showed similar binding parameters to those of the G1^{WT}-Pal-GATA interaction (Fig. 2B).

The association and dissociation kinetics of G1^{WT} binding to Pal-GATA appeared to fit the biphasic model with a high initial k_{a1} followed by an extremely low second k_{a2} . The initial k_{d1} of G1^{WT} is slightly higher than the second k_{d2} ; consequently, the initial K_D (K_{D1}) was extremely lower than the second, K_{D2} , of G1^{WT}. Similar kinetics were seen with G1^{3KA}. In contrast, k_d values of G1^{R216Q} and G1^{R216W} were increased 8- to 9-fold, while the k_a value was increased 5-fold compared to that of G1^{WT}. Consequently, in the case of G1^{R216Q} and G1^{R216W} mutants, K_D values were calculated to be higher than that of G1^{WT} (Fig. 2B). These observations showed very good coincidence with the data obtained from EMSA off-rate analyses (Fig. 2C and D). These results thus support the contention that GATA1 binds to Pal-GATA bivalently with tight and loose interactions by using the C finger and N finger, respectively (Fig. 2E).

In stark contrast, we found that the kinetic parameters of the binding of all GATA1 proteins to rPal-GATA were quite similar to those to the Single-GATA (Fig. 1C, D, and F and 3A to C). The R216 residue in the GATA1 N finger appears to recognize the orientation of palindromic sequences and stabilize the DNA-GATA1 complex on a Pal-GATA locus.

We then evaluated the association of Tandem-GATA and these

GATA mutant proteins. The analyses revealed that the SPR parameters of G1^{WT}, G1^{R216Q}, and G1^{R216W} did not fit to a 1:1 binding model, and all exhibited biphasic binding (Fig. 3D and E). Of note, R_{max} values of G1^{WT}, G1^{R216Q}, and G1^{R216W} were approximately twice as high as those for G1^{ΔNF} and G1^{3KA}, in which dimerization capacity of GATA1 was attenuated. Importantly, SPR parameters of G1^{WT}, G1^{R216Q}, and G1^{R216W} showed biphasic kinetics with extremely high initial k_{a1} and low k_{a2} irrespective of N-finger potential of DNA interaction. We surmise that the homodimers of G1^{WT}, G1^{R216Q}, and G1^{R216W} bind bivalently to Tandem-GATA via the C finger of each GATA1 molecule (Fig. 3F), while a single molecule of G1^{ΔNF} and G1^{3KA} binds to Tandem-GATA, as is the case for Single-GATA.

Verification of homodimer binding to Tandem-GATA. To define whether two molecules of GATA1 actually bind to the Tandem-GATA, we performed a sequential oligonucleotide pulldown assay utilizing both MBP-tagged and GST-tagged recombinant GATA1 proteins. To this end, we exploited biotinylated oligonucleotides containing Single-, Pal-, and Tandem-GATA, and recovered oligonucleotides were detected by streptavidin. Upon the first pulldown with anti-MBP antibody, we identified the presence of these three oligonucleotides in the precipitated aliquots, except for G1^{ΔCF} (Fig. 4A to C).

Subsequent pulldown assays using anti-GST antibody revealed

FIG 4 Sequential oligonucleotide pulldown assays. (A to F) Detection of oligo-DNA probes containing Single-GATA (A, D), Pal-GATA (B, E), and Tandem-GATA (C, F), which bind to the respective GATA1 proteins, after the first (A to C) and the second (D to F) pulldowns. Filled and open arrowheads indicate oligo-DNA probes bound to GATA1 protein and free oligo-DNA probes, respectively. P, pulldown; I, input. (G to I) Schematic diagrams of predicted binding models of GATA1 to Single-GATA (G), Pal-GATA (H), and Tandem-GATA (I).

enrichments of the oligonucleotide containing Tandem-GATA in samples of G1^{WT}, G1^{R216Q}, and G1^{R216W} but not in the other GATA1 mutant samples (Fig. 4F, arrows). In contrast, oligonucleotides containing Single- and Pal-GATA were not recovered in

precipitates of the second pulldown assay (Fig. 4D and E). These results indicate that two molecules of GATA1 tagged with MBP and GST formed dimers and bound to Tandem-GATA, while one molecule of MBP-tagged (or GST-tagged) GATA1 was bound to

FIG 5 ChPD-seq analysis to determine the locus preference of GATA1 in genomic location. (A, B) Expression levels of exogenous GATA1 proteins in MEL cells. Comparisons of expression levels of exogenous GATA1 (filled arrowhead) to lamin B (A) and endogenous GATA1 (B; open arrowhead) by the biotin-streptavidin system and immunoblot assay with an anti-GATA1 antibody, respectively. Expression ratios of exogenous to endogenous GATA1 proteins are shown in panel B. (C) Venn diagrams showing numbers and overlap of localization sites of G1^{WT} and G1^{R216Q} in MEL cells, under the conditions of relaxed (P value report threshold, 10^{-5}) and stringent (P value report threshold, 10^{-10}) MACS peak calling thresholds. (D) Table (3 by 3) showing the number distribution of ChPD-seq peaks of G1^{WT} and G1^{R216Q} sorted into high, low, and none according to the peak height. (E) Alignment of G1^{WT} and G1^{R216Q} binding profiles at *Mgst3*, *Lmo2*, and *Zfp1* genes. Green, blue, and red arrowheads indicate no-preference, G1^{WT} preference, and G1^{R216Q} preference peaks, respectively, defined by the signal intensity. Asterisks indicate GATA1-binding loci on which occupancies of G1^{WT} and G1^{R216Q} in mice were examined in ChIP-qPCR analysis shown in Fig. 13A.

FIG 6 Mutation on residue R216 specifically reduces ChPD-seq signal intensity for Pal-GATA. (A) Spreadsheet showing the details of each of ChPD-seq peaks defined in the 3 by 3 table and type of motifs contained in the peak region. The numbers of G1^{WT} preference, no-preference, and G1^{R216Q} preference peaks are highlighted. (B) Pie charts showing the rates of peaks that were defined as G1^{WT} preference, no preference, and G1^{R216Q} preference in each category of peaks containing Single-GATA alone, both Single- and Pal-GATA, or Pal-GATA alone. Numbers in parentheses denote the number of peaks. ***, $P < 0.005$.

Single- and Pal-GATA (Fig. 4G to I). Tandem-GATA was indeed recognized by the GATA1 homodimer. Thus, the binding affinity to Tandem-GATA results from the dimerization potential of GATA1 molecules.

Residue R216 contributes to GATA1 binding to Pal-GATA in a genomic location. Our experiments using SPR and EMSA demonstrate a clear correlation between residue R216 of GATA1 and GATA sequence configuration under chromatin-free conditions. To evaluate the role of residue R216 on the chromatin DNA structure, we established lines of MEL (mouse erythroleukemia) cells stably expressing G1^{WT} and G1^{R216Q} dually fused with FLAG and biotin tags by an *in vivo* biotinylation system and performed chromatin pulldown sequencing (ChPD-seq) analysis. We chose MEL clones in which almost equivalent levels of exogenous G1^{WT} and G1^{R216Q} proteins were expressed (Fig. 5A and B). Since the exogenous-to-endogenous GATA1 ratio was almost equivalent between MEL cell lines expressing G1^{WT} and G1^{R216Q}, we expected that the peak value precipitated by streptavidin would reflect the local affinity of biotinylated GATA1 to the DNA loci, in competition with endogenous GATA1. We used the parental BirA-expressing MEL clone as a negative control of ChPD-seq assay.

ChPD-seq peak data were mapped into MM9 murine genome reference. The number of mapped reads and mapping efficiency are shown in Table S1 in the supplemental material. When we used relaxed (P value report threshold, 10^{-5}) and stringent (10^{-10}) MACS peak calling thresholds, we found that more than one-half of binding sites were bound by both G1^{WT} and G1^{R216Q} and most of the remaining binding sites were preferentially bound by G1^{WT} in either case (Fig. 5C). We then divided the peaks into a 3 by 3 table according to the signal intensity and defined G1^{WT} preference peaks (depicted in blue), G1^{R216Q} preference peaks (red), and no-preference peaks (green) (Fig. 5D and E). We found that most of binding sites were enriched in the fractions of either G1^{WT} preference or no-preference peaks (Fig. 5D).

We next searched GATA motifs within the DNA sequence from the center ± 100 -bp regions and divided them into three subcategories of peaks containing “Single-GATA alone,” “both Single- and Pal-GATA,” and “Pal-GATA alone” (Fig. 6A). Since most peaks contained Single-GATA, the frequency of peaks containing the Pal-GATA alone was less than 1% of the total. In the end, we found that more than 80% of loci containing only Pal-GATA were preferentially bound by G1^{WT}, and these 153 GATA sites likely contain important Pal-GATA-like elements. In con-

FIG 7 Transient reporter/effector cotransfection/transactivation assays measuring the effect of GATA1 on endogenous Pal-GATA. (A) Schema of the reporter constructs. Pal-GATA sequences in the *Lmo2* and *Uros* gene enhancers are depicted. (B, C) Luciferase activity driven by indicated amounts of $G1^{WT}$ and $G1^{R216Q}$ for reporter constructs carrying *Lmo2* (B) and *Uros* gene enhancer (C). Fold increase in luciferase activity was determined relative to the value obtained by samples not transfected with the $G1^{WT}/G1^{R216Q}$ expression plasmid. Shown are representative data from two and three independent experiments utilizing reporter constructs containing *Lmo2* and *Uros* gene enhancers, respectively. Means $\pm$ standard deviations (SD) from triplicate cultures are shown.

trast, this $G1^{WT}$ preferential binding was significantly reduced in the loci containing Single-GATA alone and both Single- and Pal-GATA (Fig. 6B), indicating that $G1^{R216Q}$ has lower affinity for Pal-GATA. Thus, GATA1 binding to Pal-GATA is strengthened via residue R216 in the genomic landscape.

Residue R216 modifies the transactivation activity of GATA1 on Pal-GATA. To address potential roles that residue R216 plays in GATA1 transactivation activity, we performed a transient reporter/effector cotransfection/transactivation assay. For this purpose, we constructed two firefly luciferase vectors, each of which carried a Pal-GATA-containing *Lmo2* gene enhancer and a *Uros* gene enhancer, and we introduced these reporters transiently into HEK293T cells along with either $G1^{WT}$ or $G1^{R216Q}$ expression vector. However, against our expectation, $G1^{WT}$ and $G1^{R216Q}$ did not give rise to any significant difference in the luciferase activities in this system (Fig. 7), indicating that the differences in the transactivation activity of $G1^{WT}$ and $G1^{R216Q}$ are difficult to identify in this simple transfection assay system, which may not reflect the situation in native chromatin. It is likely that $G1^{R216Q}$ can activate the Pal-GATA reporter to some extent through C finger binding to one of the GATA motifs, and if the concentration of $G1^{WT}$ and $G1^{R216Q}$ in these transfections is higher than the K_D s of these proteins with Pal-GATA, the effects of a difference in affinity might not be detected.

Therefore, we made two reporter constructs in which fragments containing either Single-GATA or Pal-GATA were placed in triplicate upstream of minimal promoter-driven *secNluc* luciferase gene (Fig. 8A) and established stable HEK293T reporter cell lines. We then transfected expressing vectors for either $G1^{WT}$ or $G1^{R216Q}$ into the cells and measured luciferase activity in the medium at five time points. Since *secNluc* protein has a long half-life,

luciferase activity at the time of measurement corresponds to the accumulated amount of luciferase proteins secreted into culture medium (Fig. 8B and D).

When we used cells carrying the Single-GATA reporter, the time course of increase in luciferase activity in the culture medium of cells expressing $G1^{R216Q}$ was quite similar to that expressing $G1^{WT}$ as long as the levels of expression of GATA1 proteins were equivalent. Furthermore, the increase of the luciferase activity showed dose dependency in the expression of $G1^{R216Q}$ and $G1^{WT}$ (Fig. 8B and C), suggesting that $G1^{R216Q}$ yielded transcriptional activity almost equivalent to that of $G1^{WT}$ on Single-GATA.

In contrast and of note, when we used the cells carrying Pal-GATA reporter, the luciferase activity of $G1^{R216Q}$ -expressing cells was significantly impaired compared with that of $G1^{WT}$ at any time point and any level of GATA1 expression (Fig. 8D and E). Although the levels of luciferase activity were increased in a dose-dependent manner, the activity of $G1^{R216Q}$ never achieved that of $G1^{WT}$, indicating that a larger quantity of $G1^{R216Q}$ would be required to achieve the level of reporter activity obtained with $G1^{WT}$. Interestingly, the kinetics of luciferase values driven by $G1^{WT}$ are not influenced by the amount of $G1^{WT}$ expression vector, while those driven by $G1^{R216Q}$ were impaired especially in low-dose expression of $G1^{R216Q}$ (Fig. 8F to H). Consequently, when we compared the ratios of $G1^{R216Q}$ luciferase activity to $G1^{WT}$ luciferase activity, the activity of $G1^{R216Q}$ was much more strongly impaired in the lower expression level of GATA1 proteins or in the early time point of transfection (Fig. 8F to H). These results thus support the contention that bivalent binding on Pal-GATA via residue R216 plays an important role in enhancing transcriptional activity especially in the circumstance of low-level expression of GATA1 protein.

FIG 8 Luciferase reporter analyses reflecting the bivalent binding of GATA1 on Pal-GATA. (A) Schema of the reporter constructs. GATA motifs are depicted in pink. (B to E) Time course change of luciferase values driven by indicated amounts of G1^{WT} and G1^{R216Q} for reporter constructs carrying Single-GATA (B) and Pal-GATA (D). The 36-h point with 200 ng of G1^{WT} expression vector was set to 100 in each of the reporter cells. Means $\pm$ SD of triplicate cultures are depicted (**, $P < 0.01$). The amounts of GATA1 proteins in the cells carrying Single-GATA (C) and Pal-GATA (E) reporter constructs were estimated by immunoblotting at 36 h after the transfection of indicated amounts of expression vectors. Shown are representative data from four independent experiments. (F) Changes in luciferase values caused by the R216Q mutation for reporter construct carrying Pal-GATA. Luciferase values driven with G1^{R216Q} were divided by those with G1^{WT} at each time and each dosage point. Means $\pm$ SD for four independent experiments are depicted. *, $P < 0.05$; **, $P < 0.01$. (G) The time-dependent changes of luciferase values presented in panels B and D are plotted in linear form. The luciferase values at 36 h after the transfection of 50-ng, 100-ng, and 200-ng G1^{WT} expression vector were set to 100. Lines correspond to kinetics of luciferase values driven by G1^{WT} and G1^{R216Q}. The rates of change of luciferase values per hour are indicated. (H) The rates of change of luciferase activity on the Pal-GATA-driven reporter construct from 12 to 18 h, from 18 to 24 h, and from 24 to 36 h after the transfection of G1^{WT} or G1^{R216Q} expression vector at the indicated dosages. Means $\pm$ SD from four independent experiments are depicted. *, $P < 0.05$.

The function of residue R216 is critical for embryonic erythropoiesis. To investigate the functional defects of G1^{R216Q} mutant in mice, we exploited a transgenic complementation rescue analysis. To this end, we made an RP23-443E19-based recombinant BAC transgene that directs transgenic expression of GATA1 in mice corresponding to the entire expression profile of endoge-

nous *Gata1* gene (Fig. 9A) (31) and established three G1^{WT} (G1^{WT}-A, -B, and -C) and two G1^{R216Q} mutant (G1^{R216Q}-A and -B) transgenic mouse lines. We used a primer set designed to discriminate the transgene amplicons from those of endogenous *Gata1* by a difference in length (Fig. 9B).

We found that the expression levels of transgene-derived

FIG 9 *Gata1*-BAC transgenic mouse lines. (A) Schematic diagram of transgenic strategy. The region of *Gata1* 2nd to 5th exons in RP23-443E19 BAC clone was replaced with cDNA encoding G1^{WT} or G1^{R216Q} together with the *Neo* gene cassette, and subsequently the cassette was removed. (B) Schematic presentation of the *Gata1* transcripts from the endogenous *Gata1* and the *Gata1* transgene. Arrows indicate the positions of the primers used for RT-PCR detection in panel C to differentiate the transgene-derived *Gata1* mRNA from endogenous GATA1 transcript. A 110-bp untranslated region (UTR) was inserted just upstream of the translation start site (triangle). (C) Transgene-derived *Gata1* mRNA levels determined in three BAC-G1^{WT} (upper panels) and two BAC-G1^{R216Q} (lower panels) transgenic lines. Three transgenic mice and one nontransgenic sibling were used for each of the strains. Filled and open arrowheads indicate transgene-derived and endogenous *Gata1* transcripts, respectively. (D) Expression of transgene-derived GATA1 proteins in G1^{WT}R and R216QR mice. The expression levels of GATA1 proteins are evaluated using an anti-GATA1 antibody (upper panel). The same membrane was reblotted with an anti-lamin B antibody as a loading control (lower panel). Expression ratios of transgene-derived GATA1 to lamin B are shown at the bottom of the panel, with the ratio for the leftmost mouse set to 1.0.

GATA1 mRNAs in individual transgenic mouse lines were comparable to that of endogenous *Gata1* (Fig. 9C). We further confirmed that the level of transgene-derived GATA1 protein in fetal livers of rescued mice was nearly identical to that derived from endogenous GATA1 in the livers of wild-type mice (Fig. 9D).

We intercrossed the BAC-G1^{WT} transgenic male mice with *Gata1* knockdown heterozygous female (*Gata1*^{G1.05}/X) mice and determined the genotype of progenies. Hemizygous male (*Gata1*^{G1.05}/Y) mice died by embryonic day 12.5 (E12.5) due to dyserythropoiesis, as was the case in a previous study (2), while *Gata1*^{G1.05}/Y mice harboring BAC-G1^{WT} transgene (referred to as G1^{WT}R) were born alive at a frequency in accordance with Mendelian expectations (Fig. 10A; see also Table S2 in the supplemental material). G1^{WT}R mice were indistinguishable in appearance

from their littermates and grew to adulthood without any abnormalities in hematological indices, morphological analysis on peripheral blood smear films, and FACS analyses of fetal liver and adult bone marrow (Fig. 10B to F; see also Table S3 in the supplemental material), indicating that the phenotype caused by GATA1 deficiency is completely rescued by transgenic expression of G1^{WT} under the BAC regulation in mice.

We next evaluated the consequence of the G1^{R216Q} mutation and the importance of residue R216. To our surprise, although *Gata1*^{G1.05}/Y mice harboring BAC-G1^{R216Q} transgene (referred to as R216QR) survived beyond the critical point of *Gata1*^{G1.05}/Y embryos, they all died by E16.5, showing signs of severe anemia (Fig. 10A and 11A and B; see also Table S2 in the supplemental material). Fetal livers of R216QR appeared pale and small compared with those of wild-type littermates (Fig. 11C). The exami-

FIG 10 Analyses of G1^{WT}R mice. (A) Interbreeding results of BAC transgenic male mice with *Gata1* knockdown heterozygous female (*Gata1*^{G1.05/X}) mice. Shown are results for representative lines of mice. The numbers of dead embryos are in parentheses. Results of another study are presented in Table S3 in the supplemental material. (B, C) Macroscopic appearance of G1^{WT}R mice and their littermates at 4 (B) and 8 (C) days of age. (D) Peripheral blood smear at 6 weeks old (Wright-Giemsa staining). (E, F) FACS analyses of bone marrow cells at 6 weeks of age (E) and fetal liver cells at E12.5 (F). Dot plots for two-color staining of c-kit and CD71 (left panels) and CD71 and Ter119 (right panels) are shown.

nation of fetal liver histology revealed reduced numbers of mature enucleated red blood cells (Fig. 11D).

We found a massive reduction of CD71⁺ erythroid progenitors in R216QR fetal livers, which are a major population in fetal livers of wild-type littermates, whereas the number of CD71[−] cells was relatively maintained (Fig. 11E). Furthermore, c-kit⁺ Ter119[−] immature progenitors were accumulated in the CD71⁺ erythroid population in R216QR embryos (Fig. 11F). Thus, defects caused by R216Q mutation are wide-ranging during erythroid differentiation, from the early stage of CD71⁺ erythroid commitment to late maturation stage Ter119⁺ erythroblasts.

In contrast, the influence of the R216Q mutation in megakaryopoiesis was milder than that in erythropoiesis. An approximately 3-fold-increased frequency of cells in the megakaryocytic lineage was observed in R216QR embryos, as well as a 3-fold activation of the colony-forming capacity of megakaryocytes, probably due to the vast reduction of erythroid cells in fetal livers. Indeed, the absolute number of megakaryocyte progenitors in R216QR was approximately one-half that in wild-type embryos (Fig. 12). Taken together, these results demonstrate that the R216Q mutation of GATA1 elicits severe dyserythropoiesis and embryonic lethality in mice, suggesting that the R216Q-based im-

pairment of GATA1 binding to Pal-GATA may be involved, at least in part, in the molecular mechanisms underlying this abnormality.

Residue R216 selectively supports GATA1 binding to Pal-GATA1 in mice. To confirm the hypothesis that the R216 residue is truly responsible for the binding of GATA1 to the Pal-GATA1 in mice and responsible for the dyserythropoiesis in embryos, we finally examined the occupancy of Pal-GATA and Single-GATA by GATA1 and its mutant in several erythroid genes *in vivo*. We selected *Uros*, *Lmo2*, *Mgst3*, *Zfp1*, and *Alas2* genes and *Aqp1*, *Ppox*, *Hbb*, *Slc4a1*, *E2f4*, *Klf11*, and *Gata2* genes containing Pal-GATA and Single-GATA alone, respectively, within 300-bp proximity of the representative GATA motif. As shown in Fig. 13A, GATA1 occupancies were decreased in R216Q mutant embryos compared with those in wild-type embryos on every gene locus examined except for the *Gata2* gene. *Gata2* is expressed in stem/progenitors and downregulated by GATA1, whereas the other selected genes are implicated in the erythroid maturation. We speculate that the reduction of GATA1 occupancy on the erythroid genes may be due in part to the enlarged population of immature erythroid progenitors in the liver of R216QR embryos (Fig. 11E and F).

FIG 11 Erythroid phenotypes caused by R216Q mutation in mice. (A, B) Macroscopic appearance of $G1^{WT}$ (A) and R216QR (B) embryos at E15.5 are shown along with wild-type littermates. (C) The liver of an R216QR embryo at E13.5 is small and pale. (D) Histological analyses showing reduced number of enucleated erythrocytes in the liver of an R216QR embryo at E12.5. Hematoxylin-eosin staining. (E) Number of CD71⁺ and CD71⁻ cells in the livers of wild-type and R216QR embryos at E12.5. Means $\pm$ SD obtained from 5 wild-type and 4 R216QR embryos are depicted. **, $P < 0.01$. (F) Representative FACS analysis of wild-type and R216QR mice at E12.5. Dot plots show the c-kit and Ter119 staining for the electronically gated CD71⁺ cells. Numbers in dot plots indicate percent cells in upper left and lower right quadrants in CD71⁺ cells.

Interestingly, the difference in the ChIP signal between wild-type and R216QR embryos was increased on the loci containing Pal-GATA alone compared with those containing Single-GATA alone. We calculated the ChIP signal ratio of R216QR to that of wild-type embryos and found that the ratio on the loci containing Pal-GATA alone was significantly decreased compared to the ratio on the loci containing Single-GATA alone (Fig. 13B), supporting our contention that residue R216 specifically functions to enhance binding on Pal-GATA in mice.

DISCUSSION

Two zinc finger DNA-binding domains of GATA1 have been characterized extensively, and their links to human diseases have also been identified. This study revealed that the configurations of *cis*-acting GATA motifs located in the regulatory regions of GATA1 target genes critically influence the DNA-binding characteristics of GATA1 and the transcription of integrated reporter genes containing them. GATA1 binds monovalently to the single GATA motif, but GATA1 lacking the C finger completely loses DNA-binding ability. Of note, while three GATA1 mutants related to the N-finger functions, i.e., $G1^{R216Q}$, $G1^{3KA}$, and $G1^{4NF}$, are all able to bind to any configuration of GATA motifs in SPR analyses, the association and dissociation parameters of these mutant GATA1 proteins to Pal-GATA and Tandem-GATA are markedly different from those of the wild-type GATA1 protein. Similarly, while all three

GATA1 mutants precipitate oligonucleotides containing any configuration of GATA motifs in pulldown assays, GATA1 bivalently binds to Pal-GATA by using N plus C fingers and to Tandem-GATA by using two C fingers in a GATA1 homodimer. Thus, our present results unequivocally demonstrate that binding kinetics between GATA1 and *cis*-acting GATA motifs elicit differential transcriptional regulation.

Fine dissection of the binding modes of these GATA1 mutants revealed important differences in the selection of *cis*-acting GATA motifs. Our ChIP and ChIPD-seq analyses verify the $G1^{R216Q}$ occupancy of Pal-GATA, although the level of ChIP signal is low compared with that of $G1^{WT}$ occupancy. We interpret this result to indicate that $G1^{R216Q}$ binds to Pal-GATA in a mode similar to that of Single-GATA. In this way, $G1^{R216Q}$ is able to bind Pal-GATA in the genome and affect gene regulation, but our reporter cotransfection/transactivation assay demonstrates that this binding does not activate transcription to the level achieved by wild-type GATA1. We found that the bivalent binding of GATA1 to Pal-GATA assisted by both N and C fingers powerfully activates target genes especially in an early phase of GATA1 expression and/or in a low-dose range of GATA1 expression. Given the fact that the GATA1 expression is dramatically upregulated during erythroid differentiation (4), we propose that the bivalent binding of GATA1 to Pal-GATA induces allosteric binding effects and enables GATA1 to act for precipitous induction of the target genes

FIG 12 Influence of R216Q mutation to megakaryopoiesis. (A) Representative FACS data of fetal megakaryopoiesis at E12.5. Numbers in dot plots indicate percent cells in each gate. (B) Number of c-kit⁺ CD41⁺, c-kit⁻ CD41⁺, and c-kit⁻ CD41⁺ CD61⁺ cells in a fetal liver. Means $\pm$ SD of results obtained from 3 wild-type and 3 R216QR embryos are depicted (*, $P < 0.05$). (C) CFU-Meg assay starting with cells recovered from liver of E13.5 embryos in littermates. The average CFU-Meg counts were determined in duplicate cultures.

during the early phase of erythroid differentiation. This notion is summarized in Fig. 14. Similarly, G1^{3KA} and G1^{ANF} mutants lacking the ability to form homodimers appear to bind to Tandem-GATA as they bind to Single-GATA. We surmise that N finger of GATA1 contributes to the modulation of dynamic gene regulation during erythroid differentiation.

It has been shown that the N and C fingers of GATA1 intimately cooperate upon binding to Pal-GATA (12, 38). The N and C fingers in a single GATA1 molecule individually bind to two GATA motifs aligned in a palindromic orientation, indicating that the N finger contributes to the specific bivalent binding of GATA1 to Pal-GATA. In this regard, structural analysis of GATA3 revealed that N finger binds the opposite face of DNA that is bound by C finger and the N finger interacts with the C-terminal basic tail of the C finger inserted into the minor groove (39). This model nicely supports the notion that N finger does not contribute to GATA1 binding to rPal-GATA, in which two GATA motifs are aligned in a reverse palindromic orientation. In addition, several cases harboring substitution mutations on the R216 residue of GATA1 N finger exhibit disorders in erythroid and megakaryocytic lineages (33, 36, 40, 41). Showing very good agreement with the human case analyses, we show here that the transgenic G1^{R216Q}-rescued mice are prone to die due to defects in definitive erythropoiesis,

indicating that the contributions of the R216 residue are vital for GATA1 activity *in vivo*. We surmise, based on these observations, that GATA1 found a specific *cis*-acting target sequence, Pal-GATA, during molecular evolution, which enables efficient binding of GATA1 to DNA by exploiting the N and C fingers. This binding mode of GATA1 seems to underlie the important functions of GATA1 in humans and mice *in vivo*, which cannot be replaced by other GATA motif configurations.

One caveat for this interpretation is that the R216 residue may also harbor alternative activities that explain the *in vivo* phenotype of the transgenic G1^{R216Q} rescued mice and/or R216Q or R216W human cases. In fact, it has been reported that the R216 residue supports GATA1 binding of transcriptional coactivator LMO2 (34). It is technically difficult at present to exclude the latter possibility; nonetheless, our present results support the notion that the N and C finger binding to Pal-GATA contributes, at least in part, to the specific function of GATA1.

One of the important features of the N finger is the contribution to GATA1 homodimer formation. It has been shown that two molecules of GATA1 dimerize reciprocally via the N and C fingers (35). In this study, we found that the dimer formation is a prerequisite for two molecules of GATA1 binding to two GATA motifs aligned in a tandem orientation. Importantly, we show here that GATA1 lacking the N finger cannot dimerize. In this regard, we

FIG 13 Residue R216 has an impact on GATA1 binding to Pal-GATA in mice. (A) ChIP-qPCR of E12.5 liver cells for G1^{WT} and G1^{R216Q} using regions containing a Pal-GATA or a Single-GATA. Means $\pm$ SD of results obtained from 5 wild-type and 7 R216Q embryos are depicted. Representative data from two independent experiments are shown. *, $P < 0.05$; **, $P < 0.01$; †, $P < 0.001$; ††, $P < 0.0001$. (B) Changes in GATA1 occupancy caused by R216Q mutation. The ratios of G1^{R216Q} occupancy to G1^{WT} occupancy on the regions containing a Pal-GATA and those containing a Single-GATA except for *Gata2* gene are shown.

previously found in a transgenic complementation rescue assay that G1^{3KA}, which lacks the dimerization potential but possesses most of the other N- and C-finger functions, is unable to rescue GATA1-deficient mice from embryonic lethality even when expressed at high levels (37). These observations indicate that the homodimer formation is important to attain full GATA1 activity, and this study explains one of the molecular bases for the importance of GATA1 dimer formation.

It has been shown that the GATA1 N finger also contributes to

the GATA1 interaction with a cofactor FOG1, which has been shown to modulate the transcriptional output of GATA1 (16, 42). Several families with inherited anemia and thrombocytopenia have been reported to harbor GATA1 mutations lacking the potential to interact with FOG1 (43). In this regard, we have conducted transgenic complementation rescue analyses of GATA1-deficient mice utilizing the GATA1 V205G (G1^{V205G}) mutant that shows reduced FOG1-interacting activity and found that the rescued mice are prone to suffer from severe spherocytic hemolysis and thrombocytopenia (29, 32). In the rescued mice, expression of GATA1 target genes related to the red cell membrane structure is markedly reduced, while expression of the GATA1 target genes related to the erythroid differentiation is relatively preserved. We envisage that the association of FOG1 may also influence the GATA1 selection of specific *cis*-acting GATA motifs.

While it has been shown that numerous transcription factors create the diversity of transcriptional regulation, accumulating lines of recent evidence also support the importance of binding site configuration for generating diversity as well as for the fine-tuning of gene expression. One salient example for this emerging notion is that a pair of intimately interacting transcription factors often show orientation and distance preference in their recognition of binding motifs (44). Similarly, in the case of transcription factors working as a homodimer or a heterodimer in which each factor possesses a structurally similar DNA binding domain, their binding sites that show a high sequence similarity to each other are usually composed of two segments arranged as direct or inverted repeats. For instance, nuclear receptor family members, including retinoic acid receptor (RAR), form homo- or heterodimers to regulate a set of target genes through retinoic acid responsive elements (RAREs). Direct repeats of a hexanucleotide separated by 1 to 5 neutral nucleotides generate RAREs. Various nuclear receptor dimers preferentially bind to differently spaced elements, thereby changing their transcriptional activity (45, 46). Similarly, Maf

FIG 14 Schematic illustration for characteristic transactivation regulatory mechanism of GATA1 on Pal-GATA. The expression of GATA1 is accelerated toward erythroid maturation. The expression of target genes is conventionally influenced by the level of GATA1 that monovalently binds to Single-GATA. Bivalent binding with the N finger to Pal-GATA acts as an allosteric effector, resulting in modulation of the transcriptional activity. Consequently, GATA1 is enabled to trigger the expression of target genes more effectively in the early stage of differentiation.

family proteins harbor a highly conserved basic-region leucine zipper (bZip) domain that contributes to the specific binding to Maf recognition element (MARE). While Maf homodimers bind to MARE, heterodimers of Maf and cap-n-collar (CNC) family of bZip transcription factors, such as Nrf2, bind slightly modified *cis*-acting motifs harboring one or a few nucleotide substitutions. Of note, the difference in *cis*-acting motif recognition between Maf and CNC is created by one amino acid alteration in the bZip domain (47, 48). Our present study revealed that Pal-GATA and Tandem-GATA generate additional important regulation to GATA1 target gene expression relative to that elicited by Single-GATA.

In conclusion, we show here that GATA1 recognizes a single GATA motif or a composite of adjacent GATA motifs and exerts its diversified bindings. We also show that these binding configurations serve as one of the critical determinants of specific transcriptional regulation. This study provides a pioneering description of the configuration-specific binding of a key transcription factor and identifies new palindromic sites of potential regulatory importance for further study.

ACKNOWLEDGMENTS

We thank Aya Goto and Tatsuro Iso for technical help. We also thank Biomedical Research Core of Tohoku University Graduate School of Medicine for technical support.

We declare that we have no conflict of interest.

FUNDING INFORMATION

This work, including the efforts of Masayuki Yamamoto, was funded by Ministry of Education, Culture, Sports, Science, and Technology (MEXT) (22118002). This work, including the efforts of Cecelia Trainor, was funded by HHS | NIH | National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). This work, including the efforts of Ritsuko Shimizu, was funded by Japan Society for the Promotion of Science (JSPS) (24659142 and 15H04759). This work, including the efforts of Atsushi Hasegawa, was funded by Japan Society for the Promotion of Science (JSPS). This work, including the efforts of Masayuki Yamamoto, was funded by Naito Foundation. This work, including the efforts of Masayuki Yamamoto, was funded by Japan Agency for Medical Research and Development (AMED). This work, including the efforts of Ritsuko Shimizu, was funded by Japan Agency for Medical Research and Development (AMED). This work, including the efforts of Masayuki Yamamoto, was funded by Takeda Medical Research Foundation.

REFERENCES

- Yamamoto M, Ko LJ, Leonard MW, Beug H, Orkin SH, Engel JD. 1990. Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. *Genes Dev* 4:1650–1662. <http://dx.doi.org/10.1101/gad.4.10.1650>.
- Takahashi S, Onodera K, Motohashi H, Suwabe N, Hayashi N, Yanai N, Nabesima Y, Yamamoto M. 1997. Arrest in primitive erythroid cell development caused by promoter-specific disruption of the GATA-1 gene. *J Biol Chem* 272:12611–12615. <http://dx.doi.org/10.1074/jbc.272.19.12611>.
- Takai J, Moriguchi T, Suzuki M, Yu L, Ohneda K, Yamamoto M. 2013. The *Gata1* 5' region harbors distinct *cis*-regulatory modules that direct gene activation in erythroid cells and gene inactivation in HSCs. *Blood* 122:3450–3460. <http://dx.doi.org/10.1182/blood-2013-01-476911>.
- Kaneko H, Shimizu R, Yamamoto M. 2010. GATA factor switching during erythroid differentiation. *Curr Opin Hematol* 17:163–168. <http://dx.doi.org/10.1097/MOH.0b013e32833800b8>.
- Suzuki M, Kobayashi-Osaki M, Tsutsumi S, Pan X, Ohmori S, Takai J, Moriguchi T, Ohneda O, Ohneda K, Shimizu R, Kanki Y, Kodama T, Aburatani H, Yamamoto M. 2013. GATA factor switching from GATA2 to GATA1 contributes to erythroid differentiation. *Genes Cells* 18:921–933. <http://dx.doi.org/10.1111/gtc.12086>.
- Boyes J, Byfield P, Nakatani Y, Ogryzko V. 1998. Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature* 396:594–598. <http://dx.doi.org/10.1038/25166>.
- Collavin L, Gostissa M, Avolio F, Secco P, Ronchi A, Santoro C, Del Sal G. 2004. Modification of the erythroid transcription factor GATA-1 by SUMO-1. *Proc Natl Acad Sci U S A* 101:8870–8875. <http://dx.doi.org/10.1073/pnas.0308605101>.
- Hernandez-Hernandez A, Ray P, Litos G, Ciro M, Ottolenghi S, Beug H, Boyes J. 2006. Acetylation and MAPK phosphorylation cooperate to regulate the degradation of active GATA-1. *EMBO J* 25:3264–3274. <http://dx.doi.org/10.1038/sj.emboj.7601228>.
- Lin KR, Li CL, Yen JJ, Yang-Yen HF. 2013. Constitutive phosphorylation of GATA-1 at serine²⁶ attenuates the colony-forming activity of erythrocyte-committed progenitors. *PLoS One* 8:e64269. <http://dx.doi.org/10.1371/journal.pone.0064269>.
- Kaneko H, Kobayashi E, Yamamoto M, Shimizu R. 2012. N- and C-terminal transactivation domains of GATA1 protein coordinate hematopoietic program. *J Biol Chem* 287:21439–21449. <http://dx.doi.org/10.1074/jbc.M112.370437>.
- Trainor CD, Ghirlando R, Simpson MA. 2000. GATA zinc finger interactions modulate DNA binding and transactivation. *J Biol Chem* 275:28157–28166. <http://dx.doi.org/10.1074/jbc.M000020200>.
- Trainor CD, Omichinski JD, Vandergon TL, Gronenborn AM, Clore GM, Felsenfeld G. 1996. A palindromic regulatory site within vertebrate GATA-1 promoters requires both zinc fingers of the GATA-1 DNA-binding domain for high-affinity interaction. *Mol Cell Biol* 16:2238–2247. <http://dx.doi.org/10.1128/MCB.16.5.2238>.
- Lowry JA, Atchley WR. 2000. Molecular evolution of the GATA family of transcription factors: conservation within the DNA-binding domain. *J Mol Evol* 50:103–115.
- Shimizu R, Takahashi S, Ohneda K, Engel JD, Yamamoto M. 2001. *In vivo* requirements for GATA-1 functional domains during primitive and definitive erythropoiesis. *EMBO J* 20:5250–5260. <http://dx.doi.org/10.1093/emboj/20.18.5250>.
- Tsang AP, Visvader JE, Turner CA, Fujiwara Y, Yu C, Weiss MJ, Crossley M, Orkin SH. 1997. FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell* 90:109–119. [http://dx.doi.org/10.1016/S0092-8674\(00\)80318-9](http://dx.doi.org/10.1016/S0092-8674(00)80318-9).
- Hong W, Nakazawa M, Chen YY, Kori R, Vakoc CR, Rakowski C, Blobel GA. 2005. FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *EMBO J* 24:2367–2378. <http://dx.doi.org/10.1038/sj.emboj.7600703>.
- Johnson KD, Boyer ME, Kang JA, Wickrema A, Cantor AB, Bresnick EH. 2007. Friend of GATA-1-independent transcriptional repression: a novel mode of GATA-1 function. *Blood* 109:5230–5233. <http://dx.doi.org/10.1182/blood-2007-02-072983>.
- Letting DL, Chen YY, Rakowski C, Reedy S, Blobel GA. 2004. Context-dependent regulation of GATA-1 by friend of GATA-1. *Proc Natl Acad Sci U S A* 101:476–481. <http://dx.doi.org/10.1073/pnas.0306315101>.
- Blobel GA, Nakajima T, Eckner R, Montminy M, Orkin SH. 1998. CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proc Natl Acad Sci U S A* 95:2061–2066. <http://dx.doi.org/10.1073/pnas.95.5.2061>.
- Grass JA, Boyer ME, Pal S, Wu J, Weiss MJ, Bresnick EH. 2003. GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proc Natl Acad Sci U S A* 100:8811–8816. <http://dx.doi.org/10.1073/pnas.1432147100>.
- Stumpf M, Waskow C, Krötschel M, van Essen D, Rodriguez P, Zhang X, Guyot B, Roeder RG, Borggrete T. 2006. The mediator complex functions as a coactivator for GATA-1 in erythropoiesis via subunit Med1/TRAP220. *Proc Natl Acad Sci U S A* 103:18504–18509. <http://dx.doi.org/10.1073/pnas.0604494103>.
- Yu M, Riva L, Xie H, Schindler Y, Moran TB, Cheng Y, Yu D, Hardison R, Weiss MJ, Orkin SH, Bernstein BE, Fraenkel E, Cantor AB. 2009. Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Mol Cell* 36:682–695. <http://dx.doi.org/10.1016/j.molcel.2009.11.002>.
- Dore LC, Chlon TM, Brown CD, White KP, Crispino JD. 2012. Chromatin occupancy analysis reveals genome-wide GATA factor switching

- during hematopoiesis. *Blood* 119:3724–3733. <http://dx.doi.org/10.1182/blood-2011-09-380634>.
24. Tripic T, Deng W, Cheng Y, Zhang Y, Vakoc CR, Gregory GD, Hardison RC, Blobel GA. 2009. SCL and associated proteins distinguish active from repressive GATA transcription factor complexes. *Blood* 113:2191–2201. <http://dx.doi.org/10.1182/blood-2008-07-169417>.
 25. Wadman IA, Osada H, Grütz GG, Agulnick AD, Westphal H, Forster A, Rabbitts TH. 1997. The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J* 16:3145–3157. <http://dx.doi.org/10.1093/emboj/16.11.3145>.
 26. Ohneda K, Shimizu R, Nishimura S, Muraosa Y, Takahashi S, Engel JD, Yamamoto M. 2002. A minigene containing four discrete *cis* elements recapitulates GATA-1 gene expression *in vivo*. *Genes Cells* 7:1243–1254. <http://dx.doi.org/10.1046/j.1365-2443.2002.00595.x>.
 27. Yu L, Moriguchi T, Souma T, Takai J, Satoh H, Morito N, Engel JD, Yamamoto M. 2014. GATA2 regulates body water homeostasis through maintaining aquaporin 2 expression in renal collecting ducts. *Mol Cell Biol* 34:1929–1941. <http://dx.doi.org/10.1128/MCB.01659-13>.
 28. Lowry JA, Gamsjaeger R, Thong SY, Hung W, Kwan AH, Broitman-Maduro G, Matthews JM, Maduro M, Mackay JP. 2009. Structural analysis of MED-1 reveals unexpected diversity in the mechanism of DNA recognition by GATA-type zinc finger domains. *J Biol Chem* 284:5827–5835. <http://dx.doi.org/10.1074/jbc.M808712200>.
 29. Hasegawa A, Shimizu R, Mohandas N, Yamamoto M. 2012. Mature erythrocyte membrane homeostasis is compromised by loss of the GATA1-FOG1 interaction. *Blood* 119:2615–2623. <http://dx.doi.org/10.1182/blood-2011-09-382473>.
 30. Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27:1017–1018. <http://dx.doi.org/10.1093/bioinformatics/btr064>.
 31. Suzuki M, Moriguchi T, Ohneda K, Yamamoto M. 2009. Differential contribution of the Gata1 gene hematopoietic enhancer to erythroid differentiation. *Mol Cell Biol* 29:1163–1175. <http://dx.doi.org/10.1128/MCB.01572-08>.
 32. Shimizu R, Ohneda K, Engel JD, Trainor CD, Yamamoto M. 2004. Transgenic rescue of GATA-1-deficient mice with GATA-1 lacking a FOG-1 association site phenocopies patients with X-linked thrombocytopenia. *Blood* 103:2560–2567. <http://dx.doi.org/10.1182/blood-2003-07-2514>.
 33. Yu C, Niakan KK, Matsushita M, Stamatoyannopoulos G, Orkin SH, Raskind WH. 2002. X-linked thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting DNA binding rather than FOG-1 interaction. *Blood* 100:2040–2045. <http://dx.doi.org/10.1182/blood-2002-02-0387>.
 34. Campbell AE, Wilkinson-White L, Mackay JP, Matthews JM, Blobel GA. 2013. Analysis of disease-causing GATA1 mutations in murine gene complementation systems. *Blood* 121:5218–5227. <http://dx.doi.org/10.1182/blood-2013-03-488080>.
 35. Mackay JP, Kowalski K, Fox AH, Czolij R, King GF, Crossley M. 1998. Involvement of the N-finger in the self-association of GATA-1. *J Biol Chem* 273:30560–30567. <http://dx.doi.org/10.1074/jbc.273.46.30560>.
 36. Phillips JD, Steensma DP, Pulsipher MA, Spangrude GJ, Kushner JP. 2007. Congenital erythropoietic porphyria due to a mutation in GATA-1: the first *trans*-acting mutation causative for a human porphyria. *Blood* 109:2618–2621. <http://dx.doi.org/10.1182/blood-2006-06-022848>.
 37. Shimizu R, Trainor CD, Nishikawa K, Kobayashi M, Ohneda K, Yamamoto M. 2007. GATA-1 self-association controls erythroid development *in vivo*. *J Biol Chem* 282:15862–15871. <http://dx.doi.org/10.1074/jbc.M701936200>.
 38. Bates DL, Chen Y, Kim G, Guo L, Chen L. 2008. Crystal structures of multiple GATA zinc fingers bound to DNA reveal new insights into DNA recognition and self-association by GATA. *J Mol Biol* 381:1292–1306. <http://dx.doi.org/10.1016/j.jmb.2008.06.072>.
 39. Chen Y, Bates DL, Dey R, Chen PH, Machado AC, Laird-Offringa IA, Rohs R, Chen L. 2012. DNA binding by GATA transcription factor suggests mechanisms of DNA looping and long-range gene regulation. *Cell Rep* 2:1197–1206. <http://dx.doi.org/10.1016/j.celrep.2012.10.012>.
 40. Balduini CL, Pecci A, Loffredo G, Izzo P, Noris P, Grosso M, Bergamaschi G, Rosti V, Magrini U, Ceresa IF, Conti V, Poggi V, Savoia A. 2004. Effects of the R216Q mutation of GATA-1 on erythropoiesis and megakaryocytopoiesis. *Thromb Haemost* 91:129–140. 10.1160/TH03-05-0290.
 41. Tubman VN, Levine JE, Campagna DR, Monahan-Earley R, Dvorak AM, Neufeld EJ, Fleming MD. 2007. X-linked gray platelet syndrome due to a GATA1 Arg216Gln mutation. *Blood* 109:3297–3299. <http://dx.doi.org/10.1182/blood-2006-02-004101>.
 42. Crispino J, Lodish MB, MacKay JP, Orkin SH. 1999. Use of altered specificity mutants to probe a specific protein-protein interaction in differentiation: the GATA-1:FOG complex. *Mol Cell* 3:219–228. [http://dx.doi.org/10.1016/S1097-2765\(00\)80312-3](http://dx.doi.org/10.1016/S1097-2765(00)80312-3).
 43. Balduini CL, Savoia A. 2012. Genetics of familial forms of thrombocytopenia. *Hum Genet* 131:1821–1832. <http://dx.doi.org/10.1007/s00439-012-1215-x>.
 44. Wang J, Zhuang J, Iyer S, Lin X, Whitfield TW, Greven MC, Pierce BG, Dong X, Kundaje A, Cheng Y, Rando OJ, Birney E, Myers RM, Noble WS, Snyder M, Weng Z. 2012. Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res* 22:1798–1812. <http://dx.doi.org/10.1101/gr.139105.112>.
 45. Kurokawa R, DiRenzo J, Boehm M, Sugarman J, Gloss B, Rosenfeld MG, Heyman RA, Glass CK. 1994. Regulation of retinoid signalling by receptor polarity and allosteric control of ligand binding. *Nature* 371:528–531. <http://dx.doi.org/10.1038/371528a0>.
 46. Umesono K, Murakami KK, Thompson CC, Evans RM. 1991. Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. *Cell* 65:1255–1266. [http://dx.doi.org/10.1016/0092-8674\(91\)90020-Y](http://dx.doi.org/10.1016/0092-8674(91)90020-Y).
 47. Kimura M, Yamamoto T, Zhang J, Itoh K, Kyo M, Kamiya T, Aburatani H, Katsuoka F, Kurokawa H, Tanaka T, Motohashi H, Yamamoto M. 2007. Molecular basis distinguishing the DNA binding profile of Nrf2-Maf heterodimer from that of Maf homodimer. *J Biol Chem* 282:33681–33690. <http://dx.doi.org/10.1074/jbc.M706863200>.
 48. Yamamoto T, Kyo M, Kamiya T, Tanaka T, Engel JD, Motohashi H, Yamamoto M. 2006. Predictive base substitution rules that determine the binding and transcriptional specificity of Maf recognition elements. *Genes Cells* 11:575–591. <http://dx.doi.org/10.1111/j.1365-2443.2006.00965.x>.