Supporting Information

Wilkinson-White et al. 10.1073/pnas.1105898108

SI Methods.

Protein Production and Purification. $GATA1_{NC}$ (P17679, residues 200–318) and $GATA1_{NF}$ (P17679, residues 200–248) were expressed from pET-11A and pGEX-2T vectors, respectively, with induction by IPTG in *Escherichia coli* Rosetta 2 cells at 25 °C for 20 h.

Cells containing $GATA1_{NC}$ were lyzed by sonication in 50 mM Tris, 100 mM NaCl, 40 μM $ZnSO_4$, 0.5 mM PMSF, 0.1% (vol/vol) β-mercaptoethanol, pH 7.4, DNase (10 μg/mL), RNase (10 μg/mL) and $MgCl_2$ (10 mM) were added and the sample agitated on ice (30 min). The soluble fraction was applied to SP Sepharose Fast Flow (Sigma-Aldrich) resin and $GATA1_{NC}$ was eluted with a NaCl step gradient (0.2–1 M). The concentration of NaCl was adjusted to 0.2 M, and the protein was purified by cation exchange chromatography on a Uno S-1 column (BioRad) in 50 mM Tris, 40 μM Zn_2SO_4, 1 mM DTT pH 7.4 and eluted with 0.1–0.8 M NaCl. $GATA1_{NC}$ was further purified by size exclusion chromatography on a HiLoad 16/60 Superdex 75 column (GE Healthcare) in 50 mM Tris, 150 mM NaCl, 1 mM DTT pH 7.4.

$GATA1_{NF}$ was initially purified as previously described (1). Following this, the immobilized GST-fusion protein was eluted from glutathione (GSH) resin with 50 mM Tris, 150 mM NaCl, 10 mM $CaCl_2$, 20 mM GSH pH 7.4, and the GST removed by treatment with thrombin (100 U) overnight at RT. $GATA1_{NF}$ was then purified using cation exchange chromatography on a Uno S-1 column as for $GATA1_{NC}$. Uniformly labeled ^{15}N or $^{15}N/^{13}C$ $LMO2_{LIM2}$-$LDB1_{LID}$, $LMO2_{LIM1+2}$-$LDB1_{LID}$, $dFOG_{F1}$, $GATA1_{NC}$, and $GATA1_{NF}$ were prepared using the strategy developed by (2) and purified as described above.

Interaction Assays. ^{35}S-GATA1 constructs were produced using a Promega TnT (T7) Coupled Reticulocyte Lysate Transcription/Translation kit. Maltose binding protein (MBP)-GATA1 constructs were expressed in *E. coli* and partially purified by amylose affinity chromatography. Samples were incubated with GST-fusion proteins immobilized on GSH-sepharose beads in binding buffer (20 mM Tris, 250 mM NaCl, 0.5% (vol/vol) IGEPAL CA-630 (Sigma-Aldrich), 40 μM $ZnSO_4$, 10% (vol/vol) glycerol, 1 mM DTT, 0.2 M PMSF pH 7.6) for 1.5 h at 4 °C. After extensive washing, bound proteins were analyzed by SDS-PAGE and visualized using a PhosphorImager (Molecular Dynamics), or Western blotting using an anti-MBP monoclonal antibody (Sigma-Aldrich). All $GATA1_{NF}$ mutants showing an interaction with $LMO2_{LIM1+2}$-$LDB1_{LID}$ were purified without an affinity tag and folding was assessed by 1D ^1H NMR (Fig. S8*B*). Oligonucleotides were obtained from Sigma-Genosys, and EMSAs were carried out as described in ref. 3 except that samples were incubated without poly dIdC and separated on an 8% (wt/vol) nondenaturing polyacrylamide gel. Data were analyzed using a PhosphorImager (Molecular Dynamics).

1. Kowalski K, Czolij R, King GF, Crossley M, Mackay JP (1999) The solution structure of the N-terminal zinc finger of GATA-1 reveals a specific binding face for the transcriptional co-factor FOG. *J Biomol NMR* 13:249–262.
2. Marley J, Lu M, Brachen C (2001) A method for efficient isotopic labelling of recombinant proteins. *J Biomol NMR* 20:71–75.
3. Ryan DP, Duncan JL, Lee C, Kuchel PW, Matthews JM (2008) Assembly of the oncogenic DNA-binding complex LMO2-Ldb1-TAL1-E12. *Proteins* 70:1461–1474.

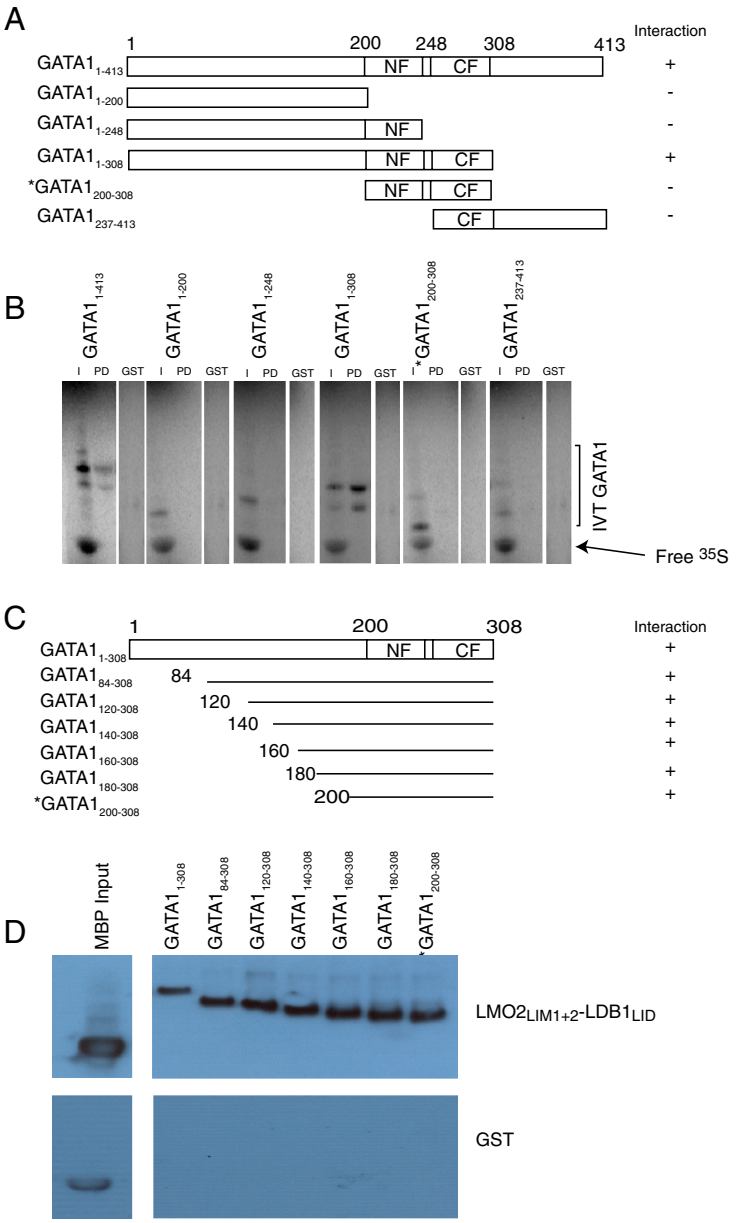

Fig. S1. *Defining the LMO2-binding domain of GATA1.* (*A*) GATA1 deletion constructs used in GST-pulldown experiments, and summary of GST-pulldown experiments. The presence of an interaction is indicated by a plus sign and no interaction by a minus sign. (*B*) GST-pulldown experiments testing the binding of in vitro transcribed/translated ^{35}S-GATA1 constructs to GST-LMO2$_{LIM1+2}$-LDB1$_{LID}$. "I" indicates 10% input of ^{35}S-GATA1 proteins, "PD" and "GST" indicate the pulldown of ^{35}S-GATA1 constructs by GST-LMO2$_{LIM1+2}$-LDB1$_{LID}$ and GST only, respectively. Signals were detected by autoradiography. (*C*) N-terminal deletion constructs of GATA1$_{1–308}$ and summary of GST-pulldown experiments. (*D*) GST-pulldown experiments (*Right*) testing the binding of MBP-GATA1 constructs to GST-LMO2$_{LIM1+2}$-LDB1$_{LID}$ (*Top*) and GST only (*Bottom*), along with an MBP control [*Left*; MBP input (*Top*) and pulldown of MBP by GST-LMO2$_{LIM1+2}$LDB1$_{LID}$ (*Bottom*)]. MBP and MBP-GATA1 were detected by Western blot using an anti-MBP antibody. GATA1$_{200–308}$, marked by an asterisk (*) is referred to as GATA1$_{NC}$ in the text and comprised both the N- and C-terminal zinc fingers of GATA1. Significant precipitation was observed in the well of the PD lane for GATA1$_{200–308}$ in *B*—a phenomenon observed for this construct only. This suggests an insoluble complex was formed under the conditions of the experiment, but a similar complex formed using the MBP-tagged GATA1 construct was soluble (*C* and *D*).

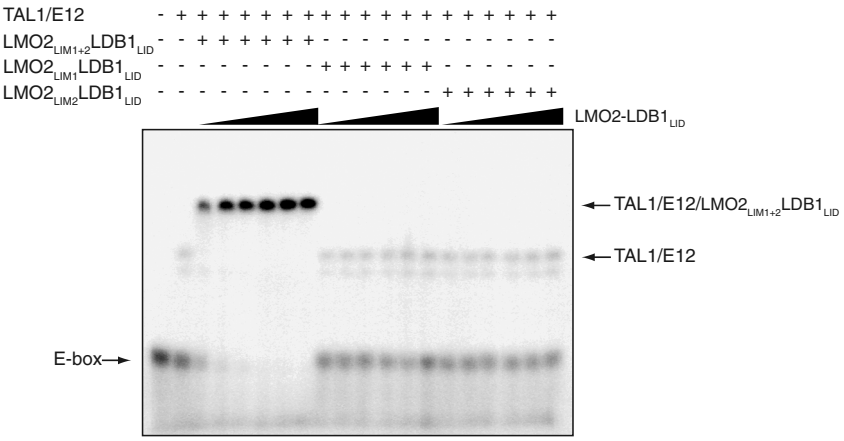

Fig. S2. *LMO2 requires both LIM domains to bind TAL1/E12 on DNA.* EMSA analysis of TAL1/E12/LMO2-LDB1$_{LID}$ complex formation on an E-box oligonucleotide, using LMO2$_{LIM1+2}$-LDB1$_{LID}$ (lanes 3–8), LMO2$_{LIM1}$-LDB1$_{LID}$ (lanes 9–14), or LMO2$_{LIM2}$-LDB1$_{LID}$ (lanes 15–20). The TAL1/E12 concentration was kept constant at 0.4 μM, and each LMO2-LDB1$_{LID}$ construct added at concentrations of 0.2, 0.4, 0.8, 1.6, 3.2, and 6.4 μM.

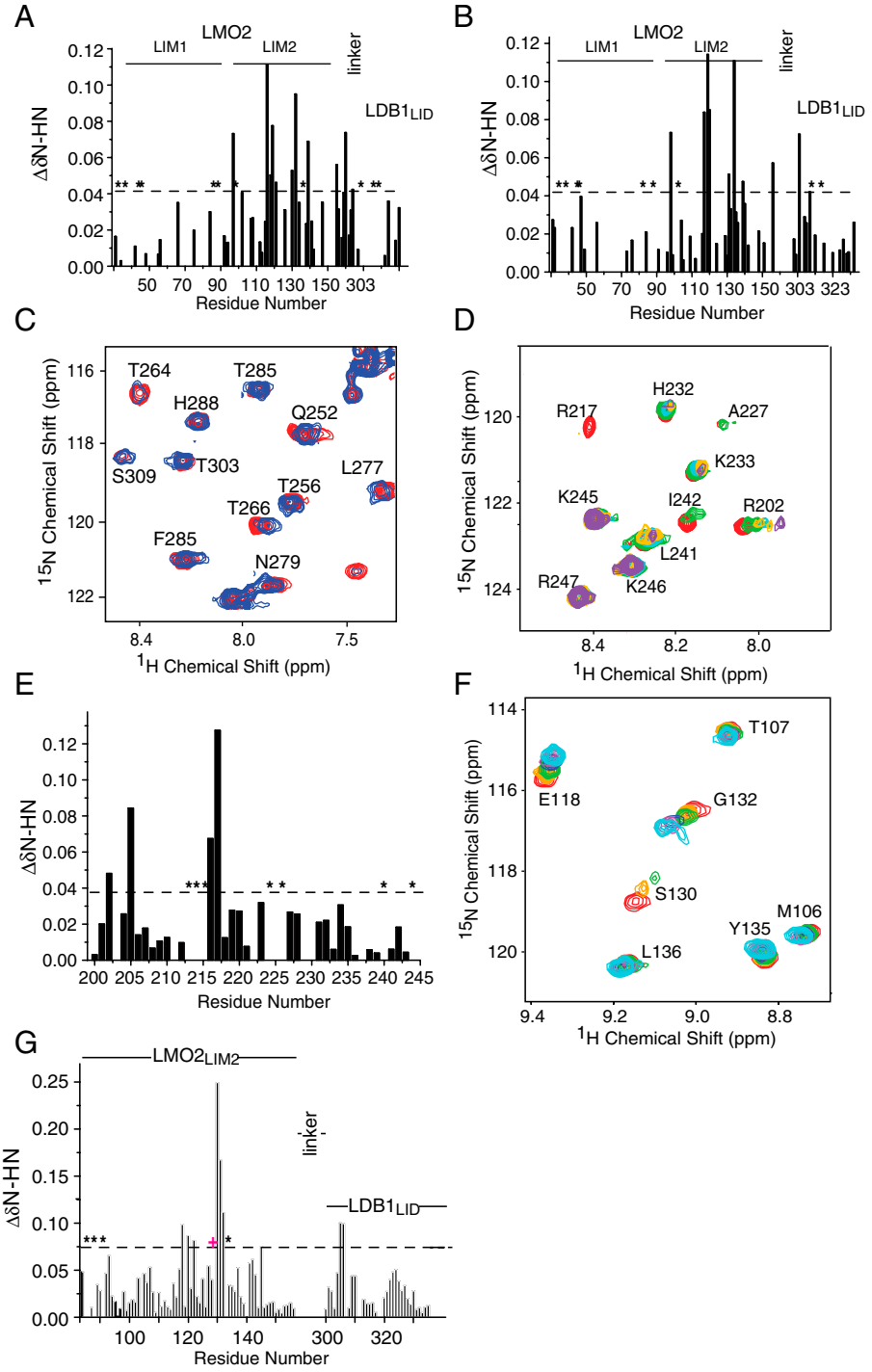

Fig. S3. *Identification of interacting domains and interface residues of GATA1 and* LMO2-LDB1$_{LID}$. Essentially identical chemical shift changes were observed for the ^{15}N-LMO2$_{LIM1+2}$-LDB1$_{LID}$ backbone amides following addition of either GATA1$_{NC}$ (*A*) or GATA1$_{NF}$ (*B*). These changes indicated that all interacting residues were from the LIM2 region of LMO2. (*C*) A portion of the ^{15}N-HSQC spectrum of ^{15}N-GATA1$_{CF}$ alone (initial concentration 200 μM, red peaks) and following addition of 2 M equivalents of LMO2$_{LIM1+2}$-LDB1$_{LID}$ (blue peaks), confirming the previous observation that all interacting residues were within the N-terminal zinc finger of GATA1. (*D*) A portion of the ^{15}N-GATA1$_{NF}$ HSQC spectrum during a titration with LMO2$_{LIM1+2}$-LDB1$_{LID}$ (red—no LMO2$_{LIM1+2}$-LDB1$_{LID}$, green—0.5 M equivalents, orange—1 M equivalent, blue—1.5 M equivalents, violet—2 M equivalents). (*E*) Weighted chemical shift changes for GATA1$_{NF}$ backbone amides from *D*. The dashed lines indicate two standard deviations above the mean chemical shift change. (*F*) A portion of the ^{15}N-HSQC spectrum of ^{15}N-LMO2$_{LIM2}$-LDB1$_{LID}$ (initial concentration 200 μM) during a titration with GATA1$_{NF}$ (red—no GATA1$_{NF}$, orange—0.25 M equivalent, green—0.5 M equivalents, blue—1 M equivalent, purple—1.5 M equivalents, cyan—2 M equivalents). (*G*) Weighted chemical shift changes for LMO2$_{LIM2}$-LDB1$_{LID}$ backbone amides and backbone amide peaks from *F*. The dashed lines indicate two standard deviations above the mean chemical shift change. For *A*, *B*, *E*, and *G*, an asterisk immediately above the dashed line indicates residues that were unable to be assigned in the free protein, with a plus sign (magenta) indicating a residue where the peak disappeared upon addition of the titrant.

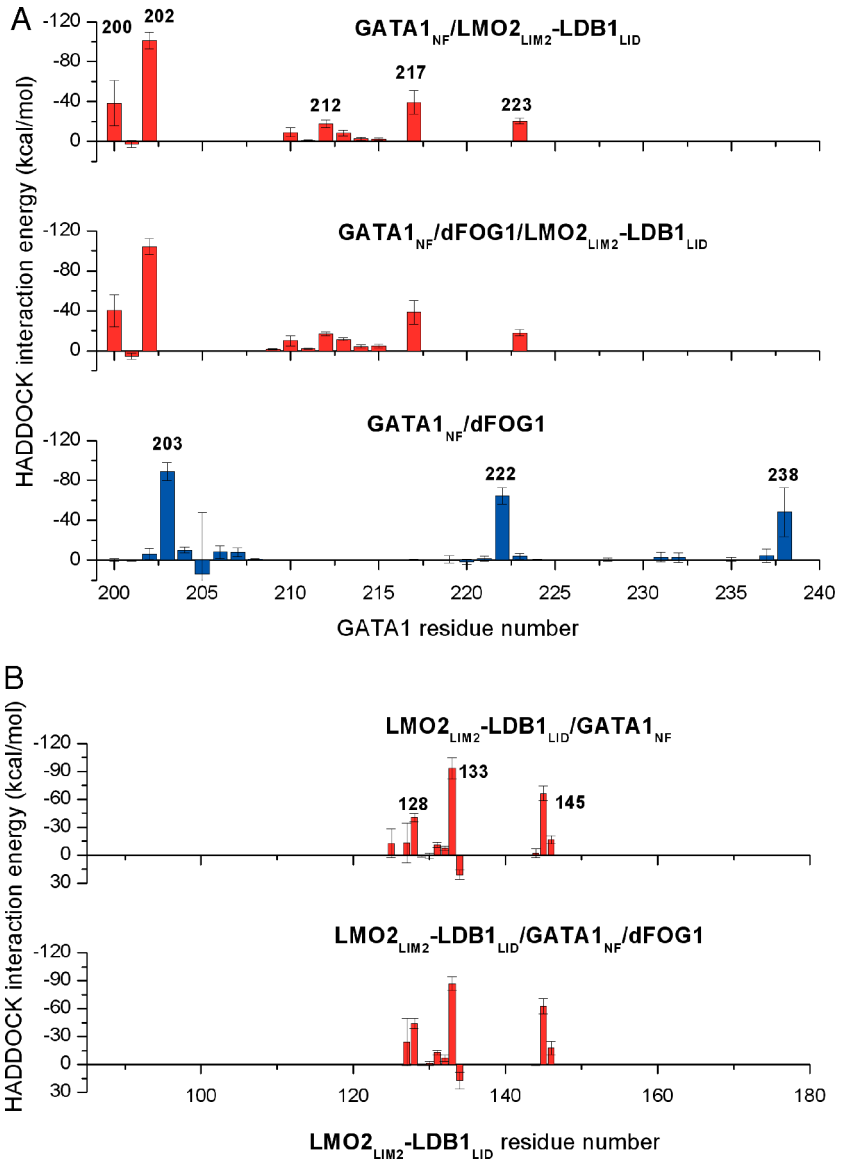

Fig. S4. *Per-residue intermolecular energies calculated in HADDOCK.* (*A*) Residues from GATA1 that contact: (*Top*) LMO2 in the $GATA1_{NF}/LMO2_{LIM2}$-$LDB1_{LID}$ model, (*Middle*) LMO2 in the $GATA1_{NF}/dFOG1_{F1}/LMO2_{LIM2}$-$LDB1_{LID}$ model, and (*Bottom*) $dFOG1_{F1}$ in the $GATA1_{NF}/dFOG1_{F1}/LMO2_{LIM2}$-$LDB1_{LID}$ model (or in the $GATA1_{NF}/dFOG1_{F1}$ structure). (*B*) Residues from LMO2 and LDB1 that contact GATA1 in the $GATA1_{NF}/LMO2_{LIM2}$-$LDB1_{LID}$ model (*Top*) and in the $GATA1_{NF}/dFOG1_{F1}/LMO2_{LIM2}$-$LDB1_{LID}$ model (*Bottom*), respectively. On the X axis the numbers correspond to LMO2, except that numbers 157–168 refer to the linker and 169–180 refer to LDB1 (300–311). In all cases, a high negative energy indicates a large amount of per residue intermolecular contact within the model. Calculated energies (±SD) of the 10 best-energy structures of each individual run are shown.

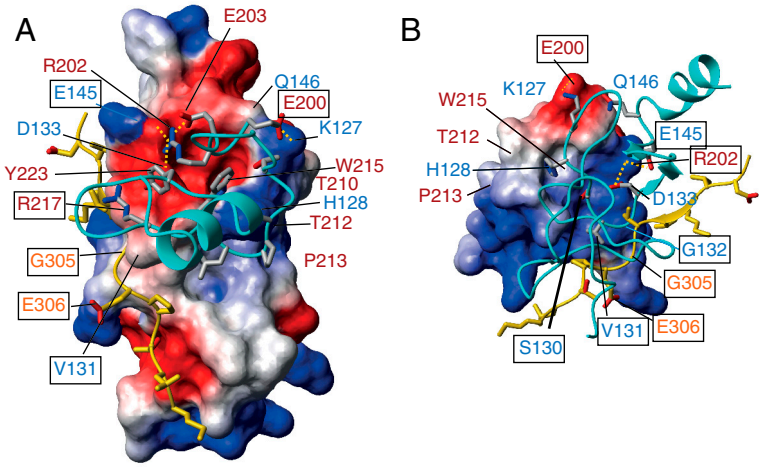

Fig. S5. (*A*) Electrostatic potential surface diagram of $LMO2_{LIM2}$ with $LDB1_{LID}$ and $GATA1_{NF}$ shown as yellow and cyan ribbons, respectively, and relevant side chains shown as sticks. (*B*) Electrostatic potential surface diagram of $GATA1_{NF}$ with LDB1 in yellow and $LMO2_{LIM2}$ in cyan and relevant side chains shown as sticks. Red is negative and blue is positive charge on the protein surface. Proposed salt bridges are shown as yellow dotted lines. Residues are labeled with blue text for LMO2, orange for LDB1, and red for GATA1. Boxed labels indicate nonburied residues from Fig. 2*D* that lie close to the interface.

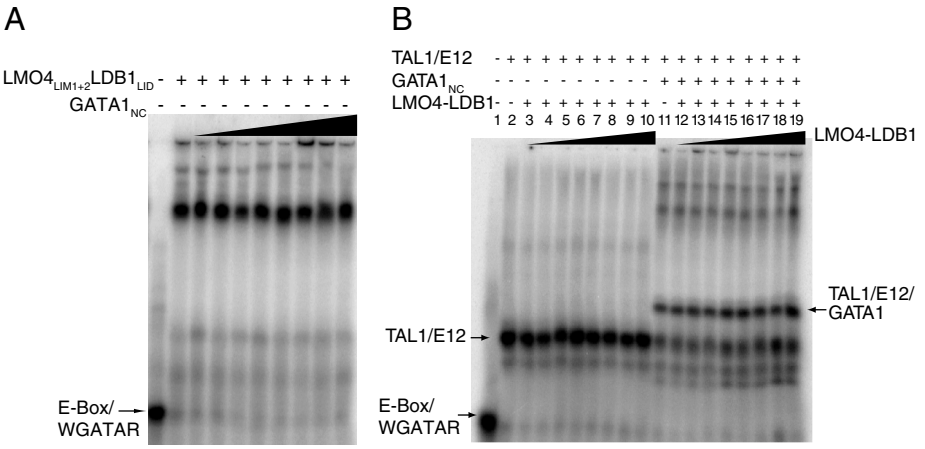

Fig. S6. *LMO4 does not bind* $GATA1_{NC}$ *or TAL1/E12 complexes.* (*A*) EMSA analysis showing $GATA1_{NC}$ bound to an E-box/WGATAR oligonucleotide does not interact with LMO4-$LDB1_{LID}$. The $GATA1_{NC}$ concentration was kept at 0.4 μM, with LMO4-$LDB1_{LID}$ added at concentrations of 0.5, 1, 2, 5, 7.5, 10, 15, and 20 μM. (*B*) EMSA analysis indicating LMO4-$LDB1_{LID}$ was also unable to bind TAL1/E12 in both the absence (lanes 3–10) and presence (lanes 11–19) of $GATA1_{NC}$. TAL1/E12 was at a concentration of 200 nM and $GATA1_{NC}$ was at 500 nM. LMO4-$LDB1_{LID}$ was added at concentrations of 5, 10, 25, 50, 100, 200, 300, and 400 nM (lanes 3–8 and 12–19).

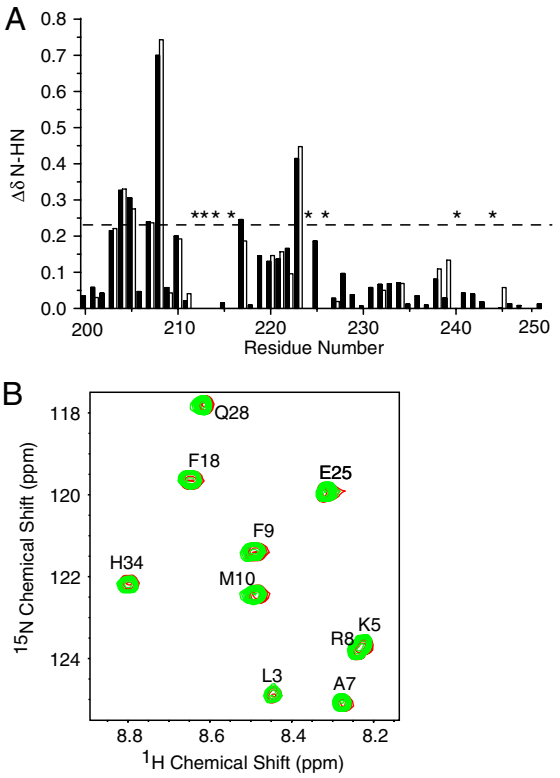

Fig. S7. $dFOG_{F1}$ *interacts with a* $GATA1_{NF}/LMO2_{LIM2}$-$LDB1_{LID}$ *complex but not* $LMO2_{LIM2}$-$LDB1_{LID}$. (*A*) Comparison of weighted chemical shift changes for backbone amides of $GATA1_{NF}$ after addition of 1.5 M equivalents dFOG1 to either ^{15}N-$GATA1_{NF}$ (black bars) or to a ^{15}N-$GATA1_{NF}/LMO2_{LIM2}$-$LDB1_{LID}$ complex (white bars). An asterisk indicates residues that could not be assigned in the spectrum of the free protein. (*B*) A portion of the ^{15}N-HSQC spectrum of ^{15}N-$dFOG_{F1}$ alone (red), and following addition of 1 M equivalent of $LMO2_{LIM2}$-$LDB1_{LID}$ (green). The initial concentration of ^{15}N-dFOG1 was 220 μM.

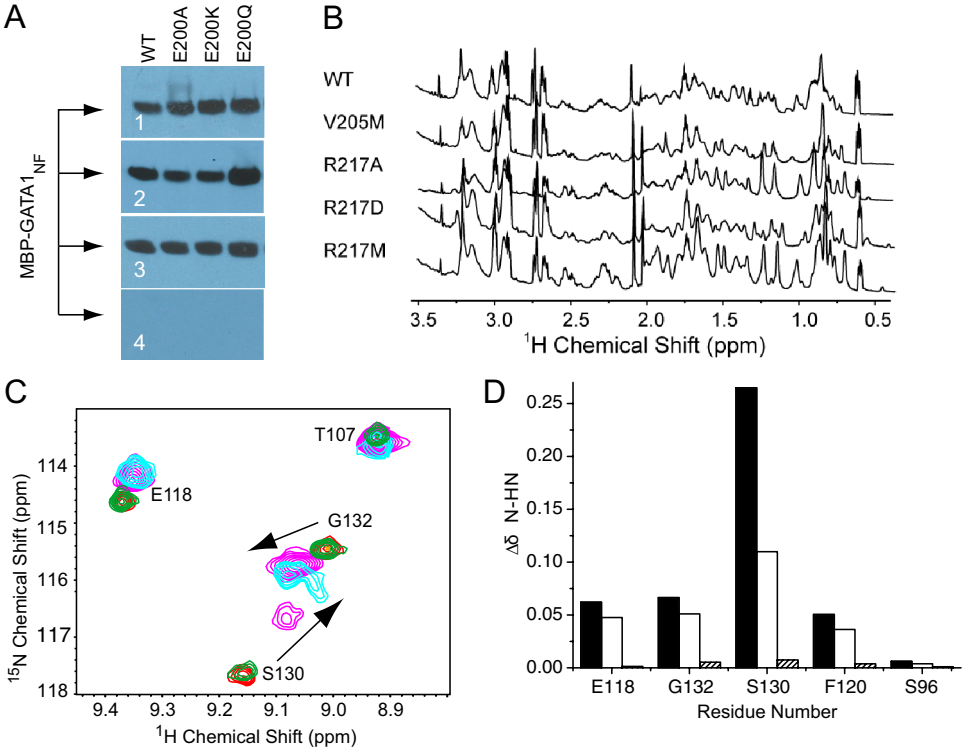

Fig. S8. Additional data for interaction of $GATA1_{NF}$ mutants to LMO2-$LDB1_{LID}$ and $dFOG_{F1}$. (*A*) GST-pulldown assays in which MBP-$GATA1_{NF}$ was pulled down by resin-bound GST-$LMO2_{LIM1+2}$-$LDB1_{LID}$ (1), GST-$LMO2_{LIM2}$-$LDB1_{LID}$ (2), GST-$dFOG_{F1}$ (3), or GST (4). Binding was detected by Western blot using anti-MBP antibodies. (*B*) ^{1}H NMR spectra (aliphatic region) for WT $GATA1_{NF}$ and point mutants that show significantly impaired binding to either LMO2-$LDB1_{LID}$ or $dFOG_{F1}$. The mutants all have sharp dispersed peaks indicating that they are folded. (*C*) A portion of the ^{15}N-HSQC spectra of ^{15}N-$LMO2_{LIM2}$-$LDB1_{LID}$ (initial concentration 220 μM; red peaks) following addition of two equivalents of either WT-$GATA1_{NF}$ (cyan), V205M-$GATA1_{NF}$ (magenta), or R217D-$GATA1_{NF}$ (green). (*D*) Weighted chemical shift changes for select residues from $LMO2_{LIM2}$-$LDB1_{LID}$ (*C*) following addition of WT-$GATA1_{NF}$ (black bars), V205M-$GATA1_{NF}$ (white bars), or R217-$GATA1_{NF}$ (hatched bars). For comparison, the residue S96 is shown, which did not undergo shifts upon addition of any $GATA1_{NF}$ species.