Supplemental Information

Analysis of disease-causing GATA1 mutations in murine gene complementation systems

Amy E. Campbell, Lorna Wilkinson-White, Joel P. Mackay, Jacqueline M. Matthews, and Gerd A. Blobel

Supplemental Methods

Plasmids

The murine GATA1 coding sequence alone or fused to the ligand binding domain of the estrogen receptor was cloned into the MSCV-based retroviral vector MIGR1 to generate MIGR1-GATA1 and MIGR1-GATA1-ER. Point mutations were introduced by site-directed mutagenesis (QuickChange, Stratagene). Murine FOG1 was cloned into MIGR1 with a single Flag tag (DYKDDDDK) fused to the C-terminus to generate MIGR1-FOG1-FLAG.

Western blot analysis

Nuclear extracts were prepared according to standard methods. Samples were fractionated through polyacrylamide gels and transferred onto nitrocellulose (Bio-Rad) or Immobilon-FL PVDF (Millipore) membranes. Blots were probed with antibodies against GATA1 (sc-265; Santa Cruz), FOG1 (sc-9361; Santa Cruz), and β-actin (A3854; Sigma-Aldrich). Signals were detected with autoradiography using chemiluminescence (Thermo Scientific) by addition of horseradish peroxidase-conjugated anti-rat antibody (A5795; Sigma-Aldrich) or with an Odyssey Imager after addition of infrared fluorescent-labeled anti-rat (926-68029 or 926-32219; LiCor) or anti-goat (926-68024 or 926-32214; LiCor) antibody. Band intensities were quantified using ImageJ.

Hemoglobin quantification

For hemoglobin assay, 5-10 x 10^6 cells were lysed in 100 μl of Drabkin's reagent (D5941; Sigma-Aldrich), and hemoglobin content was quantified by spectrophotometric

measurement of absorbance at 540 nm on a SpectraMax 190 microplate reader. Hemoglobin content was normalized to total protein, which was determined via Bradford assay.

Flow cytometry

Cells were acquired using a FACSCanto and data analyzed with BD FACSDiva software (BD Biosciences). Cell sorting was performed with a FACSDiva or FACSAria II (BD Biosciences). For G1ME lineage analysis, cells were stained in PBS with 2% FBS at room temperature for 15 minutes with the following antibodies: anti-Ter119-APC, anti-CD41-PE (BD Biosciences), and anti-CD42-PE (Emfret Analytics).

Transfections and co-immunoprecipitations

Transient transfections of HEK-293 cells were performed using polyethylenimine (PEI) at a ratio of PEI to DNA of 3:1. Cells were harvested 48 h after transfection and nuclear extracts prepared according to standard methods. Immunoprecipitation was performed with anti-FLAG beads (Sigma-Aldrich), and after extensive washing, elution was carried out with 1X FLAG peptide (Sigma-Aldrich) and bound proteins analyzed by Western blotting.

Protein production and purification

Construction, expression, and purification of $LMO2_{LIM2}$-$Ldb1_{LID}$, dFOG-F1 and GATA1 NF have been previously described.[1-3] GATA1 NF mutants were expressed and purified

using the same strategy as wild-type GATA1, with one-dimensional ^{1}H NMR used to confirm protein folding.

Isothermal titration calorimetry

Experiments were performed using a VP-ITC microcalorimeter (MicroCal) at 20 °C in 10 mM sodium phosphate, 10 mM NaCl, 1 mM DTT pH 7.5. For FOG/NF titrations, dFOG-F1 (200 μM) was titrated into a solution containing GATA1 (19 μM), G208S (20 μM), G208R (18 μM), R216W (18 μM), or D218Y (19 μM).[4] For NF/DNA titrations, GATA1 (189 μM), G208R (196 μM), R216W (187 μM) or D218Y (215 μM) were titrated into a solution containing a double-stranded 16-bp oligonucleotide containing a GATC site (19-30 μM).[4] Data were analyzed using Origin 5.0 (MicroCal).

NMR spectroscopy

Chemical shift perturbation experiments were carried out using ^{15}N-$LMO2_{LIM2}$-$Ldb1_{LID}$ (200 μM) in 20 mM Mes, 100 mM NaCl, 1 mM DTT pH 6.5. Increasing amounts of unlabelled GATA1 NF (wild-type or mutant) were added up to 1 molar equivalent, with ^{1}H, ^{15}N HSQC spectra recorded after each addition. The weighted average chemical shift changes for the peaks E118, G132 and V131 were calculated using the equation $\Delta avg = (((\delta H)^2 + (\delta N \times 0.154)^2)/2)^{1/2}$.[5]

Primers

Mutagenesis primers (5' to 3')
GATA1_V205M
Forward: GGCCAGAGAGTGTATGAACTGTGGAGCAACG
Reverse: CGTTGCTCCACAGTTCATACACTCTCTGGCC

GATA1_G208S
Forward: GAGAGTGTGTGAACTGTTCAGCAACGGCTACTCC
Reverse: GGAGTAGCCGTTGCTGAACAGTTCACACACTCTC

GATA1_G208R
Forward: GAGAGTGTGTGAACTGTAGAGCAACGGCTACTCC
Reverse: GGAGTAGCCGTTGCTCTACAGTTCACACACTCTC

GATA1_R216Q
Forward: GGCTACTCCACTGTGGCAGAGGGACAGGACAGG
Reverse: CCTGTCCTGTCCCTCTGCCACAGTGGAGTAGCC

GATA1_R216W
Forward: GGCTACTCCACTGTGGTGGAGGGACAGGACAGG
Reverse: CCTGTCCTGTCCCTCCACCACAGTGGAGTAGCC

GATA1_D218G
Forward: CCACTGTGGCGGAGGGGCAGGACAGGTCACTACC
Reverse: GGTAGTGACCTGTCCTGCCCCTCCGCCACAGTGG

GATA1_D218Y
Forward: CCACTGTGGCGGAGGTACAGGACAGGTCACTACC
Reverse: GGTAGTGACCTGTCCTGTACCTCCGCCACAGTGG

GATA1_R202Q
Forward: GCCCCTTGTGAGGCCCAAGAGTGTGTGAACTGTGG
Reverse: CCACAGTTCACACACTCTTGGGCCTCACAAGGGGC

GATA1_R217M
Forward: ACTCCACTGTGGCGGATGGACAGGACAGGTCACTACC
Reverse: GGTAGTGACCTGTCCTGTCCATCCGCCACAGTGGAGT

Transcript qPCR primers (5' to 3')
Alad
Forward: ACTTTCACCCACTGCTTCGGA
Reverse: CGATAGGCTGGACATCATCAGG

Alas2
Forward: TATGTGCAGGCCATCAACTACCCA
Reverse: TTTCCATCATCTGAGGGCTGTGGT

Alox5
Forward: ACTACATCTACCTCAGCCTCATT
Reverse: GGTGACATCGTAGGAGTCCAC

Alox12
Forward: CACACATGGTGAGGAAATGG
Reverse: GATCACTGAAGTGGGGCTGT

Casp3
Forward: TGGCAACGGAATTCGAGTCCTTCT
Reverse: TGAGCATGGACACAATACACGGGA

Clec4d
Forward: CGAGAGTAACGTGCATCCGA
Reverse: AACAGGACAGCAGGTCCAAGTAC

Cmtm7
Forward: GTGCGACCTGATAATGATCCTC
Reverse: GGTGAGCACACGGTAGAAGC

Cpox
Forward: ATGCAATTTGAAGCCAGTCCGTGG
Reverse: TCCCTTTGTAGAATAGCCTGGCCT

Eraf
Forward: GCCATGACAGAATTCCAGCAA
Reverse: TTTGGACTTCAGAAAGGTCCTGTAT

Fech
Forward: AGAGAAGCGAGGTGGTCATTCTGT
Reverse: ATGACTTTGTGGACAGTGGCTCCT

Fyb
Forward: CGAAGTTCAACACGGGGAGTA
Reverse: TGTATTCCAGAAGGCGAGCTT

Gata2
Forward: CACCCCTAAGCAGAGAAGCAA
Reverse: TGGCACCACAGTTGACACACT

Gcnt2
Forward: CTACGCGGGAAAGTTTTCGC
Reverse: GTAGAGGTTGGGCAGGCTTA

Gp6
Forward: GTCTCTCCCAAGTGACCAGCTT
Reverse: GAGCTGAGAGTGAGGGTTTAGCA

Gypa
Forward: TCACACGGCCCCTACTGAAGTGT

Reverse: TCCCTGCCATCACGCGGAAAAT

Hba-a1
Forward: GTGGATCCCGTCAACTTCAAG
Reverse: CAAGGTCACCAGCAGGCAGT

Hbb-b1
Forward: AACGATGGCCTGAATCACTTG
Reverse: AGCCTGAAGTTCTCAGGATCC

Hmbs
Forward: ATTCCAAGAGGAGCCCAGCTAG
Reverse: CATCATTAAGCTGCCGTGCA

Kit
Forward: AGCAGATCTCGGACAGCACC
Reverse: TGCAGTTTGCCAAGTTGGAG

Lrrc39
Forward: AAACACCATGAAGACGTGAAGCGG
Reverse: TTCCATCTTCCCTGGTCACCTTCT

Ly75
Forward: CACGGACAACCTGCGAATGT
Reverse: GTATTGGCTACGGCATATCCATC

Myb
Forward: TGACTTTCGACACATGGCTCCTCA
Reverse: AATGCACTTGGTGCTGCTCTCAAC

Ms4a2
Forward: TGGTTGGTTTGATATGCCTTTGT
Reverse: CACTGCACCCCAGAATGGATA

Ppox
Forward: TGCTGTAGAATTCTGGCCTCCGAA
Reverse: TCCATCCTGTCCGTGTGCAGATAA

Reep6
Forward: GAAACACCCTCACAGCAGCACAAA
Reverse: TGCACACCTTGGAGGAGAGATGTT

Rragd
Forward: TTTGGGACTCTTTGCTGCCAGTTG
Reverse: TCTTATTGCTCTGGCCTGGAGCTT

Slc4a1
Forward: TGGAGGCCTGATCCGTGATA
Reverse: AGCGCATCGGTGATGTCA

Urod
Forward: CCTACATTCGTGATGTGGCCA
Reverse: GGCAAAATGTCCATCCTTAGCA

Uros
Forward: ATGGAACGCAGATTCGGAGA
Reverse: TGACCTAATGGCCAGTGAACCT

Vwf
Forward: TCATCGCTCCAGCCACATTCCATA
Reverse: AGCCACGCTCACAGTGGTTATACA

Zfpm1
Forward: CCTTGCTACCGCAGTCATCA
Reverse: ACCAGATCCCGCAGTCTTTG

ChIP qPCR primers (5' to 3')
Abcb10 -6 kb
Forward: AGTCTCTATGCCCAGCTTCTTTGG
Reverse: AAGACAACGAGAGGAACAGGCAGT

Alad +4 kb
Forward: ACTGTGCAGCCTTCCCTCAAGTAA
Reverse: TAGCCACAAGATAAGCCAGTGCCA

Alas2 +2 kb
Forward: AGGGCAGGACTTTGCCTCTAATCT
Reverse: AGATGTCCCAGTTCCTGCAGGTTT

Gata2 -3.9 kb
Forward: GAGATGAGCTAATCCCGCTGTA
Reverse: AAGGCTGTATTTTTCCAGGCC

Hbb HS2
Forward: GGGTGTGTGGCCAGATGTTT
Reverse: CACCTTCCCTGTGGACTTCCT

Hbb HS3
Forward: CTAGGGACTGAGAGAGGCTGCTT
Reverse: ATGGGACCTCTGATAGACACATCT

Hba-a1 -12 kb

Forward: AACCCTGACTCAAAACAACAAAGTAA
Reverse: GGTTTCTGAGTTTCCTTATCTGCAA

Hbb-b1 prom
Forward: CAGGGAGAAATATGCTTGTCATCA
Reverse: GTGAGCAGATTGGCCCTTACC

Il9r prom
Forward: ACATTGCCGAGGACACAGTTCTCT
Reverse: TGCATTGCGGAAGGTGAGTCTGTA

Kit +33 kb
Forward: TGGCAGTCCTGGTTGTAGCA
Reverse: GCTGCAAGCATGCGATCA

Lrrc39 prom
Forward: TTCCCTGGTGTCTGTAGGAACACA
Reverse: GGGCTTCTGTGCAAAGGTTCAACT

Lyl1 prom
Forward: TCAGCATTGCTTCTTATCAGCC
Reverse: CGCAGAGGCCAGAGGATG

Pkhd1l1 +0.4 kb
Forward: CCGTTCTTCTTGCTCTCCTTGTGT
Reverse: AGCTTACCCTGGAAGTGACAGACA

Rragd +4 kb
Forward: CTTGGAATCCGAGGAAATGA
Reverse: TGGATATCCTCTGGGGAGTG

Slc4a1 +1.6 kb
Forward: ATCAGAAGCAACCTAGAGTCCAGC
Reverse: TAAGAGTGTAGGACCAGCAGGCAA

Slc4a1 prom
Forward: CTGAGCAGTCAAGCCTTAGTTCAC
Reverse: CCTGTCCAGTCCCTAAGGTCTTT

Slc22a4 +1.2 kb
Forward: CAGCAATGGTGGGAAGGCAGATTT
Reverse: ACAAGCACTGTTTCTGGCAAGAGC

Spna1 prom
Forward: ATGCCTCACTTTGTCCTGAGCTGT
Reverse: TGTCCCTCTGGGCCTTGTTTATCA

Tph1 prom
Forward: ACACGCTTTCACAGAAACCACCAC
Reverse: TCTCCCATTAACCGCGTTTCCTCT

Tubb1 +3 kb
Forward: CTGTGTTGACTTGAAGGCCTTTGG
Reverse: TGACTCCTGTGGCACATAAGGGTA

Uros +6 kb
Forward: TACTGCCTGGAAAGGCAAGTGAGA
Reverse: ATCTGCCACTGGTATGTCCCAGAA

Uros +11 kb
Forward: ACTGCCTTGGTGCTCGTGTGATAA
Reverse: AGCTTGCAGAACTTAGCAGCTTCC

Zfpm1 +2 kb
Forward: CTTTTCTCCTGCCCAGTCG
Reverse: TGCTGTTGCCTCGAACC

Supplemental References

1. Liew CK, Kowalski K, Fox AH, et al. Solution structures of two CCHC zinc fingers from the FOG family protein U-shaped that mediate protein-protein interactions. *Structure*. 2000;8(11):1157–1166.
2. Ryan DP, Sunde M, Kwan AHY, et al. Identification of the key LMO2-binding determinants on Ldb1. *J Mol Biol*. 2006;359(1):66–75.
3. Wilkinson-White L, Gamsjaeger R, Dastmalchi S, et al. Structural basis of simultaneous recruitment of the transcriptional regulators LMO2 and FOG1/ZFPM1 by the transcription factor GATA1. *Proc Natl Acad Sci USA*. 2011;108(35):14443–14448.
4. Liew CK, Simpson RJY, Kwan AHY, et al. Zinc fingers as protein recognition motifs: structural basis for the GATA-1/friend of GATA interaction. *Proc Natl Acad Sci USA*. 2005;102(3):583–588.
5. Ayed A, Mulder FA, Yi GS, et al. Latent and active p53 are identical in conformation. *Nat. Struct. Biol.* 2001;8(9):756–760.

Table S1. Hematologic profiles of patients carrying GATA1 mutations

Mutation	Hgb (g/dL) (10.5 - 16)	Hct (%) (33 - 46)	RBC (x10^6/µL) (3.9 - 5.2)	MCV (µm^3) (70 - 100)	MCH (pg) (23 - 34)	Retic (%) (0.5 - 1.5)	nRBC (%) (0)	RDW (11.5 - 14.5)	Plt (x10^3/µL) (150 - 400)	MPV (x10^{-15} L) (8 - 12)	Tpo (pg/mL) (60 - 150)
V205M n = 2	8.4	28.3	3.1	92.4	27.5	3	27	20.6	17.5	-	-
G208S n = 3	-	43	-	100.4	31.9	-	-	-	24	10.5	262.7
G208R n = 3	9.5	30	2.97	104	-	4	-	17.8	13.3	-	1,755
R216Q n = 7	11	42.6	5.4	78	-	-	-	18.3	73.5	9.5	-
R216W n = 1	7.0	-	-	67	<20	5	-	-	70	-	-
D218G n = 4	13.4	42	4.5	-	-	-	-	15.3	34	-	-
D218Y n = 1	9.7	33.5	3.4	-	-	-	-	17.2	12	-	-

Normal values for each parameter are indicated in parentheses. The number of individuals with reported hematologic indices is noted below each mutation. Indices shown are the average of all individuals. Hgb, hemoglobin concentration; Hct, hematocrit; RCB, red blood cell number; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; Retic, reticulocytes as percentage of total red blood cells; nRBC, number of nucleated red blood cells per 100 white blood cells; RDW, red blood cell distribution width; Plt, platelet number; MPV, mean platelet volume; Tpo, thrombopoietin concentration. Data were gathered from [27-32, 34-38].

Table S2. Examined erythroid GATA1 target genes display similar expression profiles in response to GATA1 mutations

Sensitive to GATA1 mutations	
Hbb-b1	*Gypa*
Hba-a1	*Hri*
Abcb10	*Epb4.9*
Eraf	*Ank1*
Alas2	*Spna1*
Alad	*Spnb1*
Hmbs	*Gstt1*
Uros	*Cetn4*
Urod	*Slc4a1**
Cpox	*Lrrc39**
Ppox	*Rragd**
Fech	*Slc22a4**
Myb	*Myc*
Kit	*Hhex*
Gata2†	*Casp3*
Insensitive to GATA1 mutations	
Zfpm1	*Klf1*
Lyl1	*Adamts5*
Clec4d	*Rgs18*
Tmem44	*Car1*

*More sensitive to R216Q and D218G
†Less sensitive to G208S

Table S3. Examined megakaryocyte GATA1 target genes display similar expression profiles in response to GATA1 mutations

Sensitive to GATA1 mutations	
Vwf	*Pf4*
Alox12	*Thbs1*
Capn2	*Itgb3*
Ppbp	*Gp6*
Tubb1	*Itga2b*
Gata2	*Slc22a3*
Il1rl1	*Fcgr3*
Ly75	*Cmtm7*
Insensitive to GATA1 mutations	
Asb1	*Pkn3*
Reep6	*Blvrb*
Gcnt2	*Clec4e*
Kit	*Hmga2*

Table S4. FOG and DNA binding affinities of GATA1 mutants

Protein	k_a	
	GATA1:FOG, M^{-1}	GATA1:DNA, M^{-1}
GATA1	$(4.5 \pm 1.8) \times 10^4$*	$(1.9 \pm 0.1) \times 10^6$*
	$(6.6 \pm 0.1) \times 10^4$	$(1.6 \pm 0.3) \times 10^7$
V205M	No binding*	$(2.0 \pm 1.6) \times 10^7$*
G208S	$(1.3 \pm 0.7) \times 10^5$†	$(9.1 \pm 5.2) \times 10^6$*
G208R	No binding	$(1.8 \pm 0.3) \times 10^7$
R216Q	$(9.4 \pm 0.3) \times 10^4$*	No binding*
R216W	$(1.7 \pm 0.1) \times 10^5$	No binding
D218G	$(1.9 \pm 0.3) \times 10^4$*	$>10^7$*
D218Y	No binding	$(2.4 \pm 1.0) \times 10^7$

Association constants are the average of two independent experiments. The error reports the range associated with that mean.
*Reproduced from [26].
†Previously published data indicated no detectable interaction.[26] New experiments reveal an interaction with wild-type like affinity but low enthalpy.

Table S5. Examined GATA1 occupied elements

FOG1-dependent, single GATA motif	
Hbb-b1 prom	*Hba-a1* -12 kb
Slc4a1 prom	*Rragd* +4 kb
FOG1-dependent, palindromic GATA motif	
Slc4a1 +1.6 kb	*Slc22a4* +1.2 kb
Alad +4 kb	
FOG1-independent, single GATA motif	
Hbb HS3	*Hbb* HS2
Eraf prom	*Klf1* prom
Gata2 -3.9 kb	*Lrrc39* prom
FOG1-independent, palindromic GATA motif	
Lyl1 prom	*Zfpm1* +2 kb
Spna1 prom	*Alas2* +2 kb
Abcb10 -6 kb	*Kit* +33 kb
Uros +11 kb	*Uros* +6 kb

Table S6. Summary of disease-associated GATA1 mutations

GATA1 mutation	Associated human disease	Erythroid (G1E) maturation	MK (G1ME) maturation	DNA binding		FOG1 binding		TAL1 complex binding		Proposed mechanism of disease
				In vitro	In vivo	In vitro	In vivo	In vitro	In vivo	
Wild-type		+++	+++	+++	+++	+++	+++	+++	+++	
V205M	Severe macrothrombocytopenia Severe dyserythropoietic anemia	–	–	+++	+*	–	–	+++	+++	Reduced FOG1 binding
G208S	Macrothrombocytopenia Mild dyserythropoiesis without anemia	+	–	+++	++*	+	+	ND	+++	Reduced FOG1 binding
G208R	Severe macrothrombocytopenia Severe dyserythropoietic anemia	–	–	+++	+*	–	–	ND	+++	Reduced FOG1 binding
R216Q	Macrothrombocytopenia Mild dyserythropoiesis without anemia Mild B-thalassemia	++	+	–	+++	+++	+++	–	+	Reduced TAL1 complex binding
R216W	Thrombocytopenia Congenital erythropoietic porphyria	++	++	–	+++	+++	+++	+++	+++	?
D218G	Macrothrombocytopenia Mild dyserythropoiesis without anemia	++	+	+++	+++	+++	+++	+	+	Reduced TAL1 complex binding
D218Y	Severe macrothrombocytopenia Severe dyserythropoietic anemia	–	–	+++	+*	–	–	+++	+++	Reduced FOG1 binding

*The reduced DNA binding observed intracellularly is context-dependent and a secondary consequence of impaired FOG1 binding, not a direct effect on the DNA interaction surface. MK, megakaryocyte; ND, not determined.

Supplemental Figure Legends

Figure S1. G1E cells expressing GATA1 mutants. (A and B) Anti-GATA1 and, as loading control, anti-β-actin western blots of extracts from G1E cells expressing indicated versions of GATA1 after 24 h of E2 treatment. MEL cell extracts served as a control. Bands were quantified by densitometry and relative GATA1 levels are noted below each panel. ND = not determined. Black vertical line demarcates an unrelated excised lane. (C) Representative images of cell pellets (top) and hemoglobin concentrations (bottom) after 72 h of E2 treatment. (D) One-dimensional ^{1}H-NMR spectra of GATA1 NF proteins. (E and F) Expression of GATA1-regulated genes after 24 h and 48 h of E2 treatment as determined by RT-qPCR, normalized to β-actin and plotted as fold change from uninfected samples. All error bars denote SEM (n = 3).

Figure S2. Effect of GATA1 mutations on genes encoding heme biosynthetic enzymes. (A) Expression of all eight heme biosynthetic enzymes in G1E cells expressing indicated GATA1 versions after 24 h and 48 h of E2 treatment as determined by RT-qPCR, normalized to β-actin and plotted as fold change from uninfected samples. (B) Results from (A) at 24 h were plotted as fold change from GATA1-expressing cells. (C) Expression of additional genes implicated in heme biosynthesis after 24 h of E2 treatment as determined by microarray. Data are shown as fold change from wild-type. All error bars denote SEM (n = 3).

Figure S3. G1ME cells expressing GATA1 mutants. (A) Western blot showing GATA1 expression in G1ME lines 72 h after infection with wild-type or mutant proteins

and compared to endogenous levels in MEL cells. β-actin serves as a loading control. Bands were quantified by densitometry and relative GATA1 levels are noted below each panel. (B and C) Surface expression of the terminal megakaryocyte marker CD42 and the erythroid marker Ter119 as assessed by flow cytometry in G1ME cells 72 h after infection with wild-type or mutant GATA1. Dot plots show a representative experiment. Percentages refer to the fraction of GFP positive cells expressing CD42 or Ter119. Approximately 45% of transduced cells were GFP positive. Error bars denote SEM (n = 3).

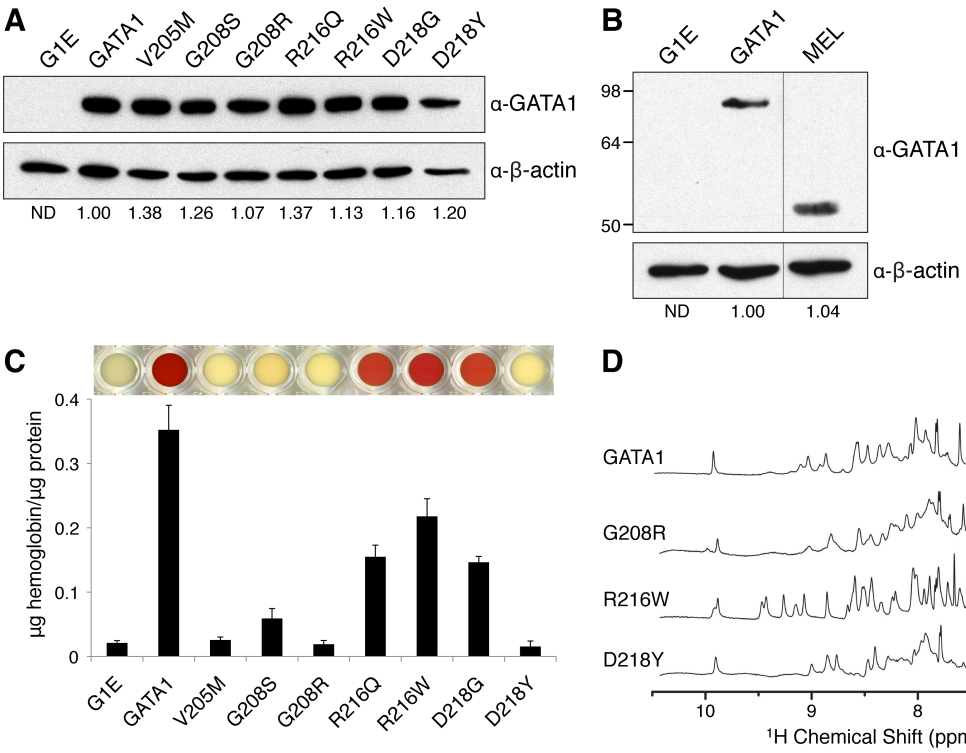
Figure S4. FOG1 binding properties of GATA1 mutants. (A) ITC data showing titration of the first ZF of the *Drosophila* FOG-family protein (dFOG-F1) into wild-type or mutant GATA1 NF. (B) Anti-GATA1 and anti-FOG1 ChIP in G1E cells expressing wild-type or mutant GATA1 after 24 h of E2 treatment using primers as in Figure 3. FOG1 ChIP signals were normalized to GATA1 ChIP signals at each site. (C) Expression of mast cell-specific genes after 24 h of E2 treatment as determined by RT-qPCR, normalized to β-actin and plotted as fold change from uninfected samples. All error bars denote SEM (n = 3).

Figure S5. TAL1 complex binding properties of GATA1 mutants. (A, B, and C) ChIP with indicated antibodies in G1E cells expressing wild-type or mutant GATA1 after 24 h of E2 treatment using primers as in Figure 3. TAL1 ChIP signals were normalized to GATA1 ChIP signals at each site. Error bars denote SEM. N = 3 for GATA1 and LMO2 ChIP, N = 2 for TAL1 ChIP.

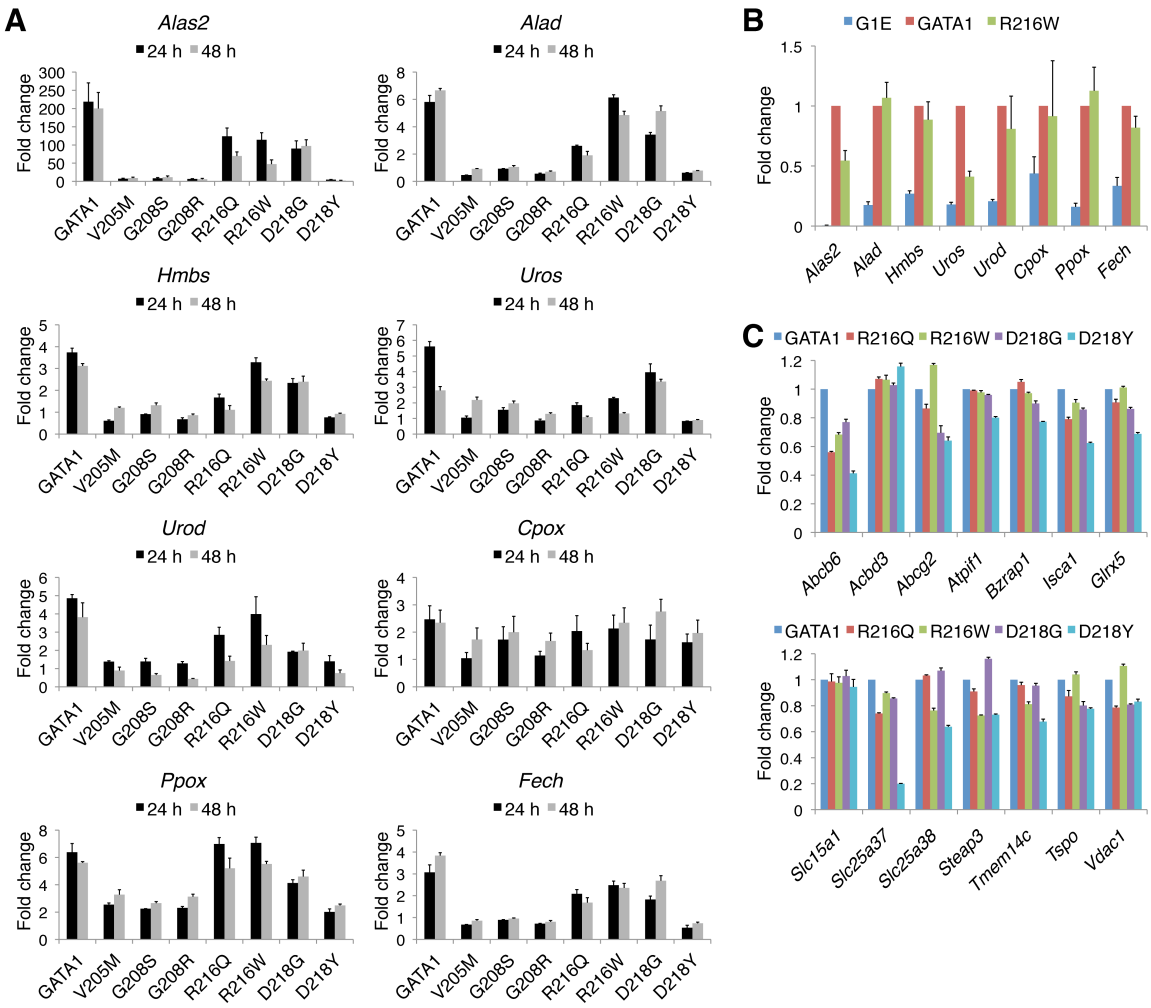
Figure S6. Engineered LMO2-disrupting mutations and factor occupancy at misregulated genes. (A) Western blot showing GATA1 expression in G1E cells expressing wild-type or mutant GATA1 after 24 h of E2 treatment. β-actin serves as a loading control. Bands were quantified by densitometry and relative GATA1 levels are noted below each panel. ND = not determined. (B) May-Grünwald-Giemsa (MGG) and benzidine staining after 72 h of E2 treatment. The percentage of hemoglobin-positive cells is indicated in the upper right corner of each benzidine panel. Scale bars, 20 μm (left panels) and 50 μm (right panels). Compare to R216Q and D218G mutants in Figure 1B. (C) Hemoglobin concentrations and representative images of cell pellets after 72 h of E2 treatment. Compare to R216Q and D218G mutants in Figure S1C. (D) Expression of GATA1-regulated genes after 24 h of E2 treatment as determined by RT-qPCR, normalized to β-actin and plotted as fold change from GATA1-expressing cells. (E and F) Anti-GATA1, FOG1, and LMO2 ChIP after 24 h of E2 treatment. FOG1 and LMO2 ChIP signals were normalized to GATA1 ChIP signals at this binding site. (G) Anti-GATA1 ChIP after 24 h of E2 treatment using primers recognizing binding sites at genes whose activities are at least 2-fold increased or decreased when compared to wild-type GATA1 expressing cells. All error bars denote SEM (n = 3).

Figure S7. Possible physical causes of differential effects of R216 and D218 mutations. GATA1 NF (grey; surface representation), FOG1 (cyan; surface representation), LMO2-Ldb1 (orange; surface representation). (A-C) R216 mutations. Wild-type R216 shown in red, Q216 shown in purple and W216 shown in magenta. Whereas the R216W mutation is similar in size to the wild-type side chain, and may be

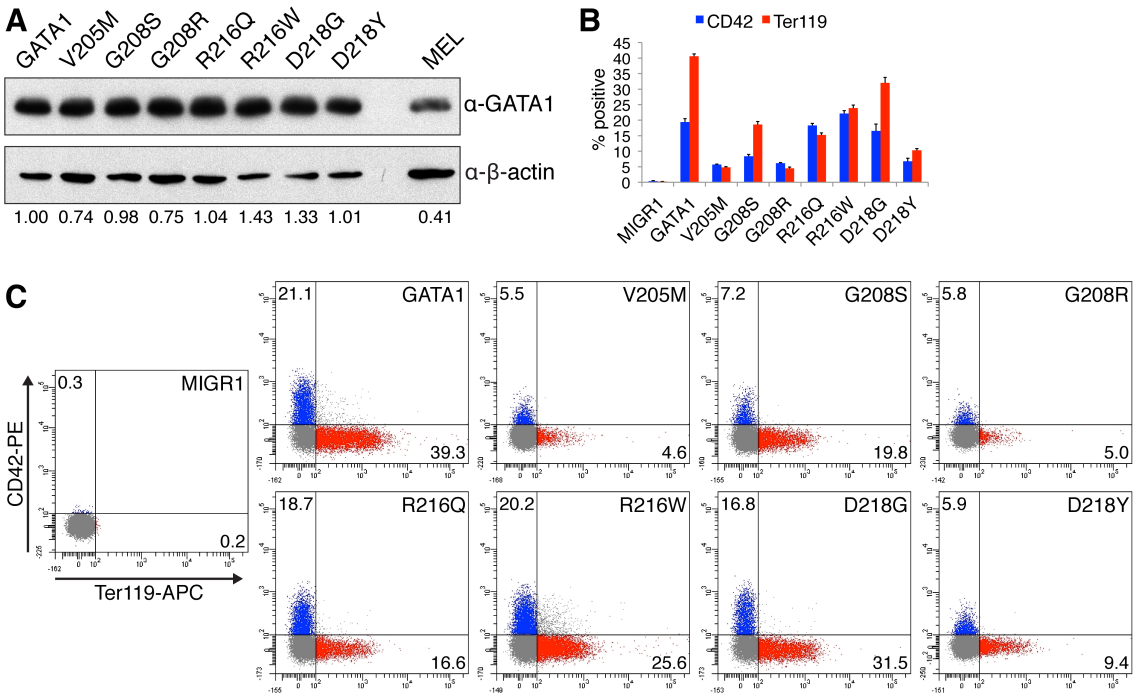
stabilized by various new interactions as indicated (dashed lines), the R216Q mutation is smaller leading to possible collapse of the LMO2 binding face (black arrow) due to repacking of side chains. (D-F) D218 mutations. Wild-type D218 (red), G218 (purple) and Y218 in magenta. D218G truncates the side chain and introduces flexibility into the backbone, leading to possible collapse of the LMO2 binding face (black arrow) due to repacking of the side chains. The larger D218Y mutation could maintain the position of the LMO2 binding face, but require repacking of the small hydrophobic core of the NF (black arrow) leading to disruption of the FOG1 binding face. Models are based on the coordinates of PDB accession code 2L6Z.



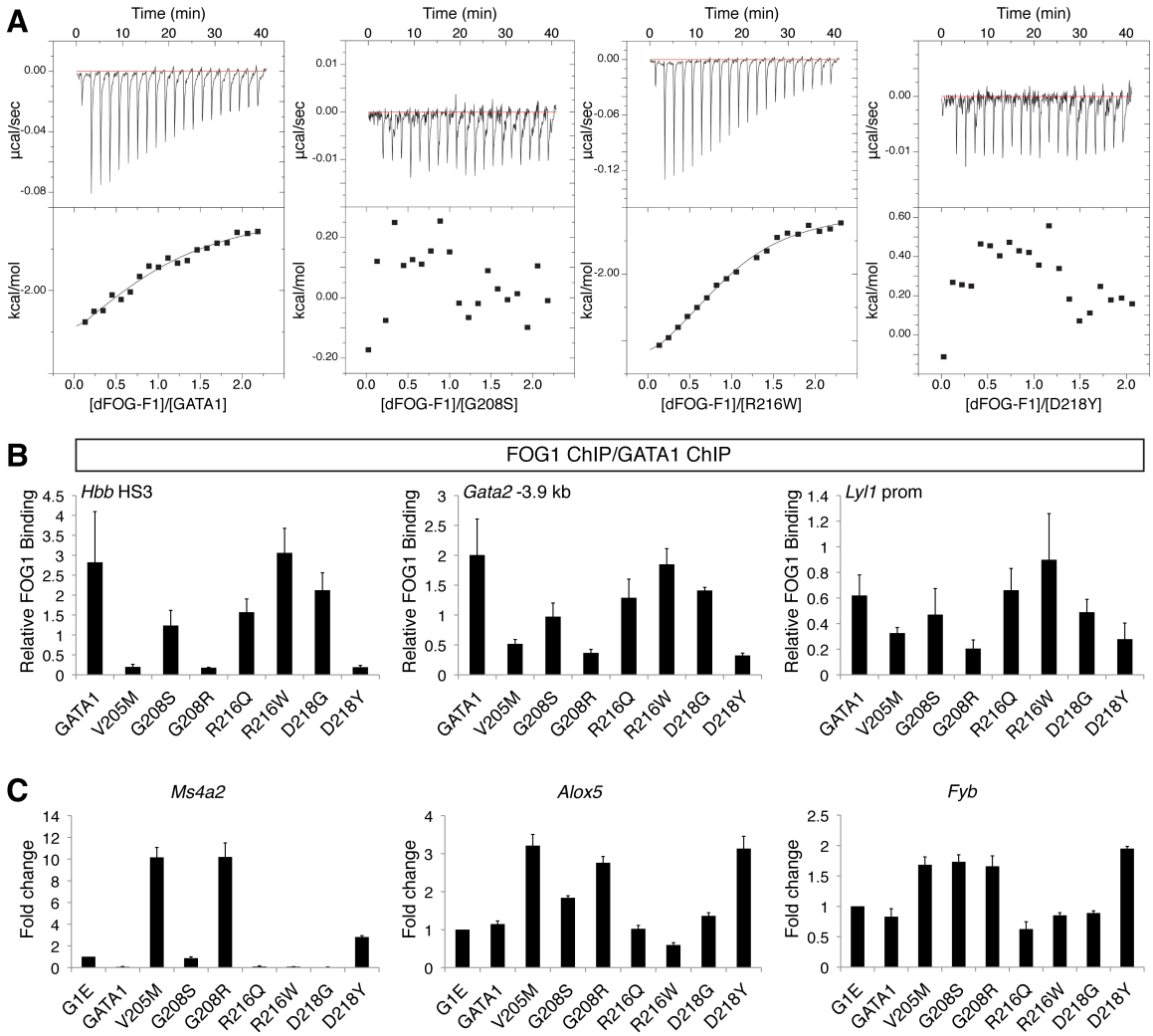
Supplemental Figure 1



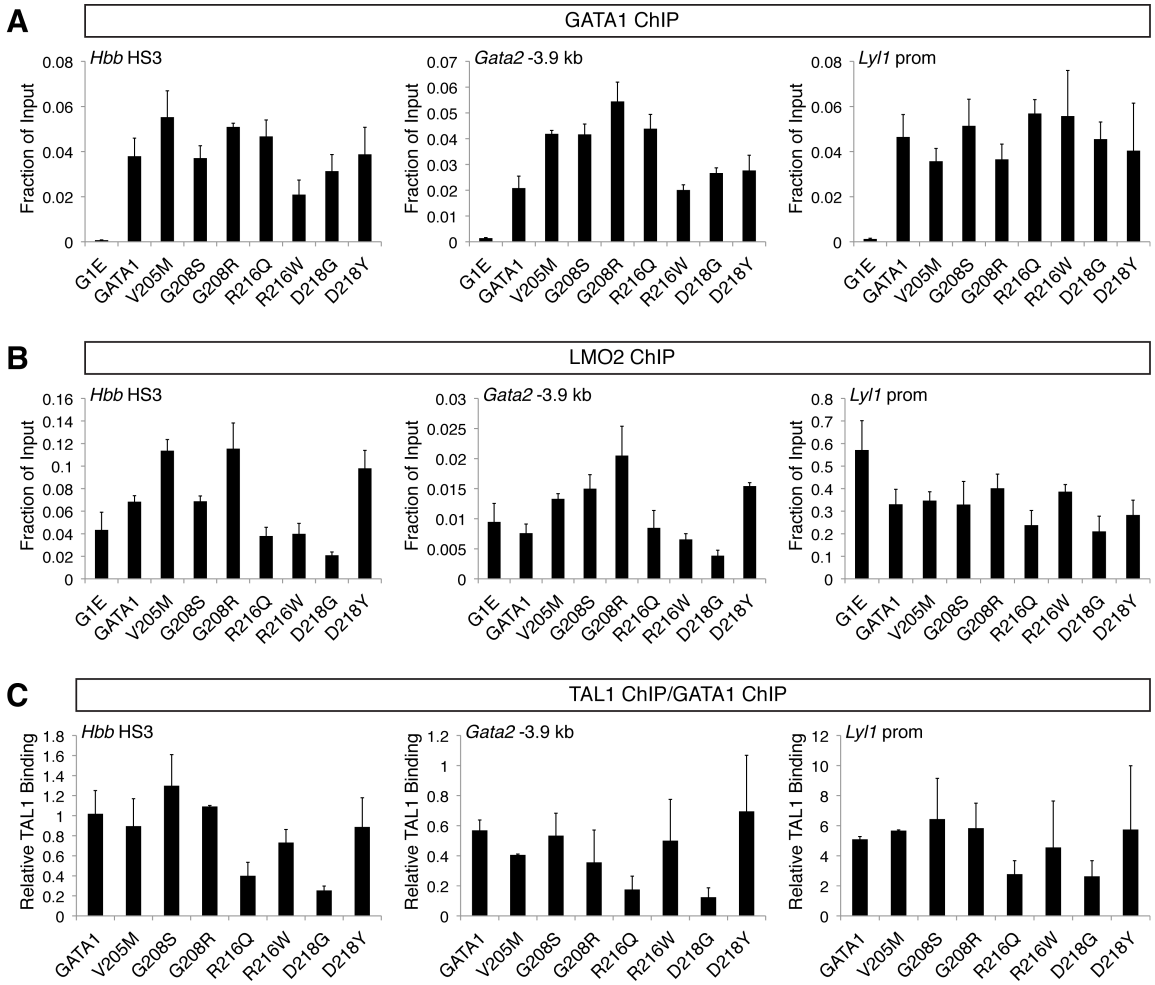
Supplemental Figure 2



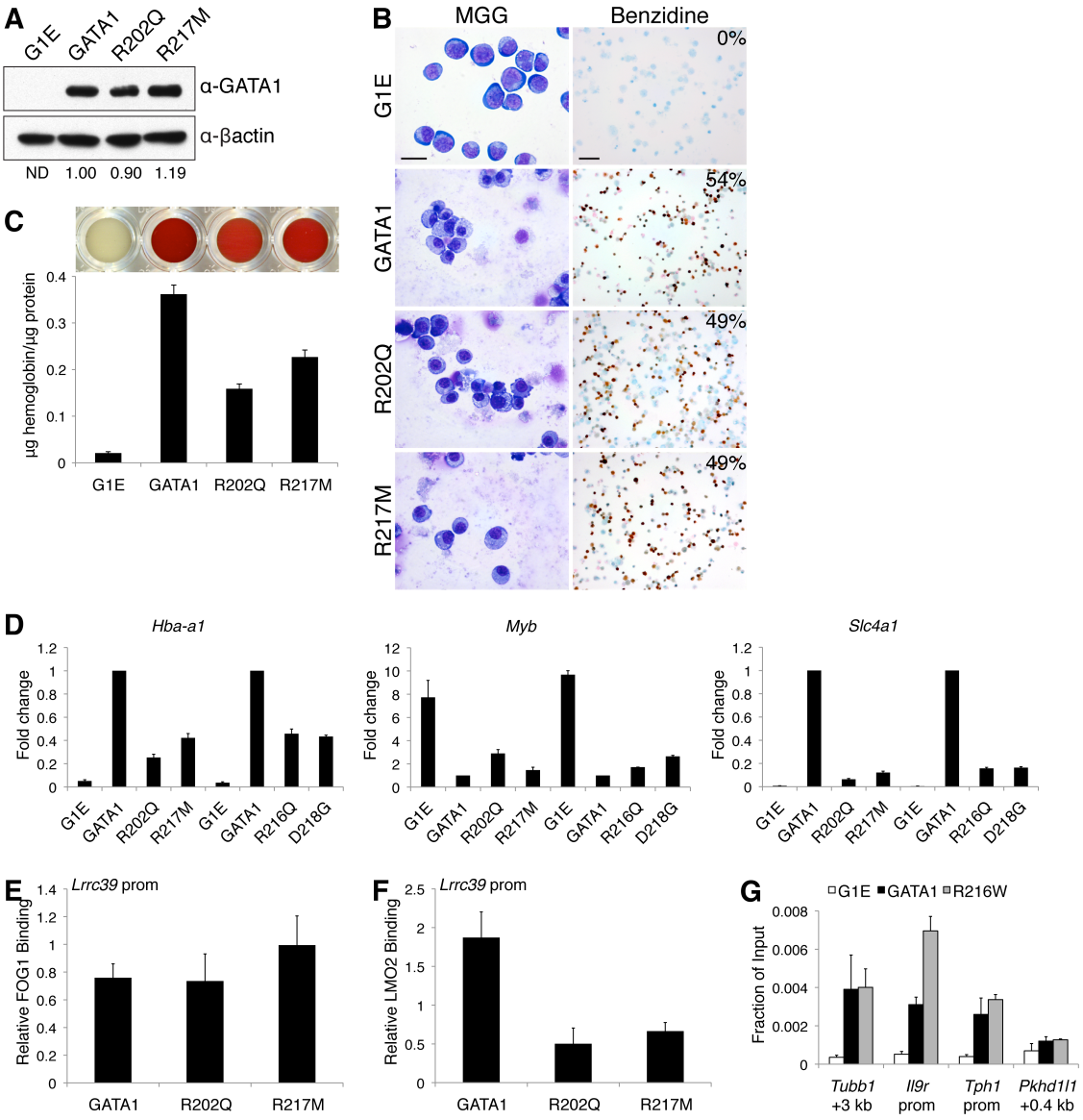
Supplemental Figure 3



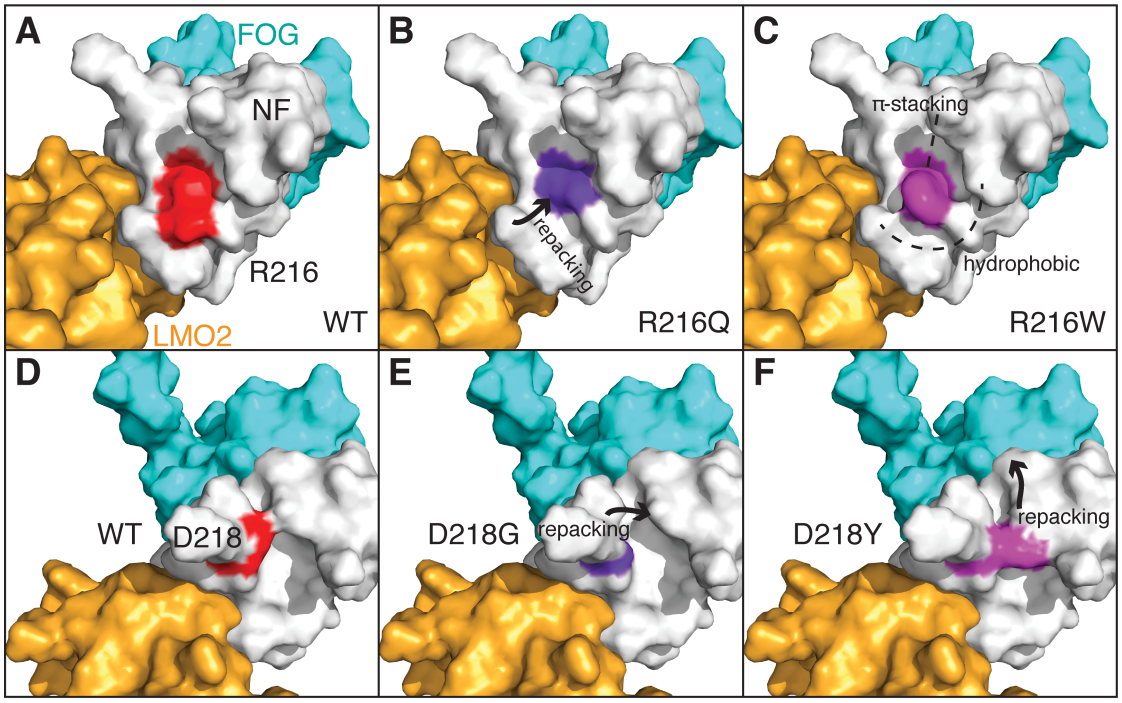
Supplemental Figure 4



Supplemental Figure 5



Supplemental Figure 6



Supplemental Figure 7