Supplementary Data

This document includes:

Supplementary Figures and legends 1 to 7

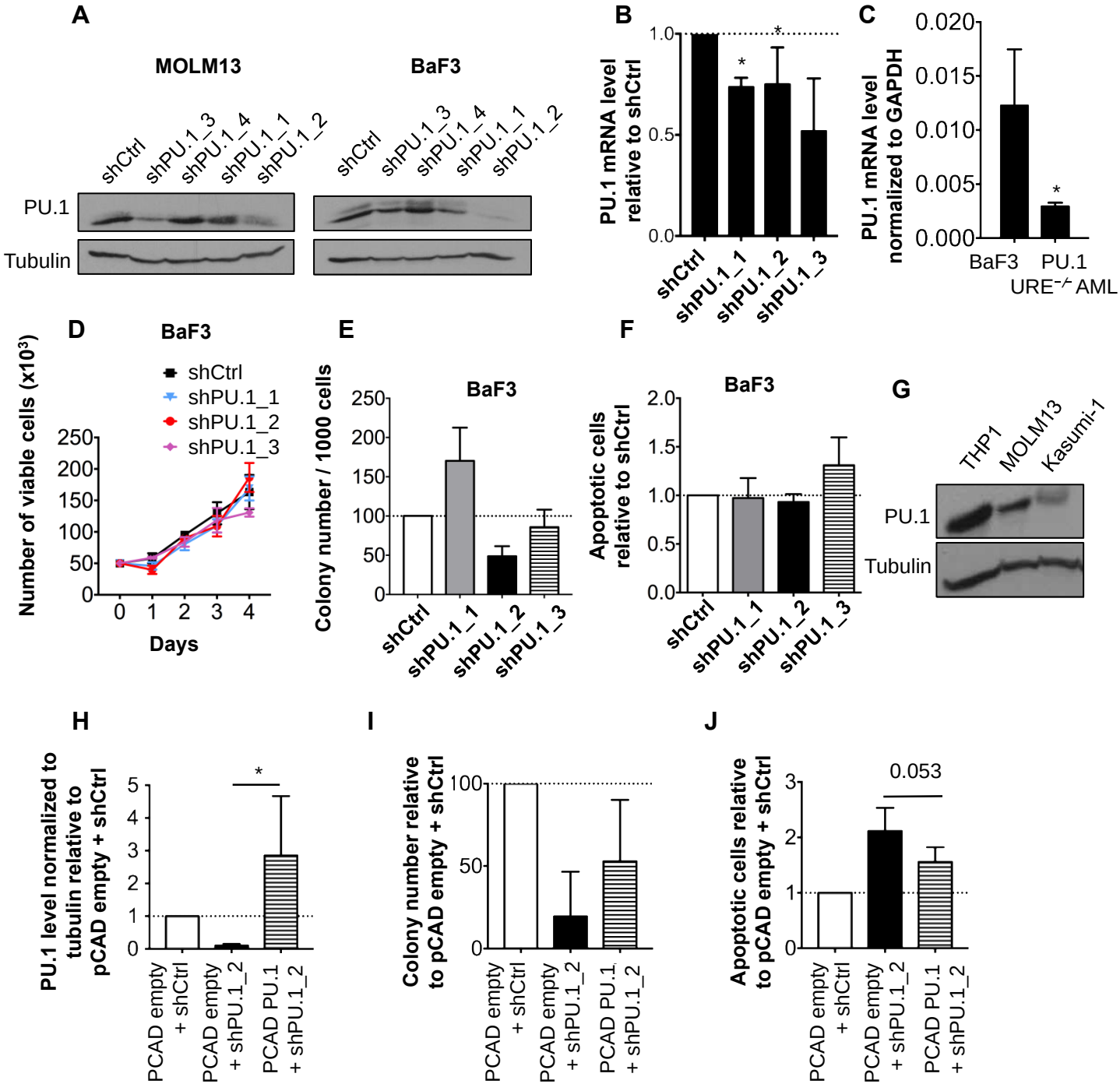
Supplementary Tables 1 to 3

Supplementary Methods

Supplementary References

SUPPLEMENTARY FIGURES AND LEGENDS

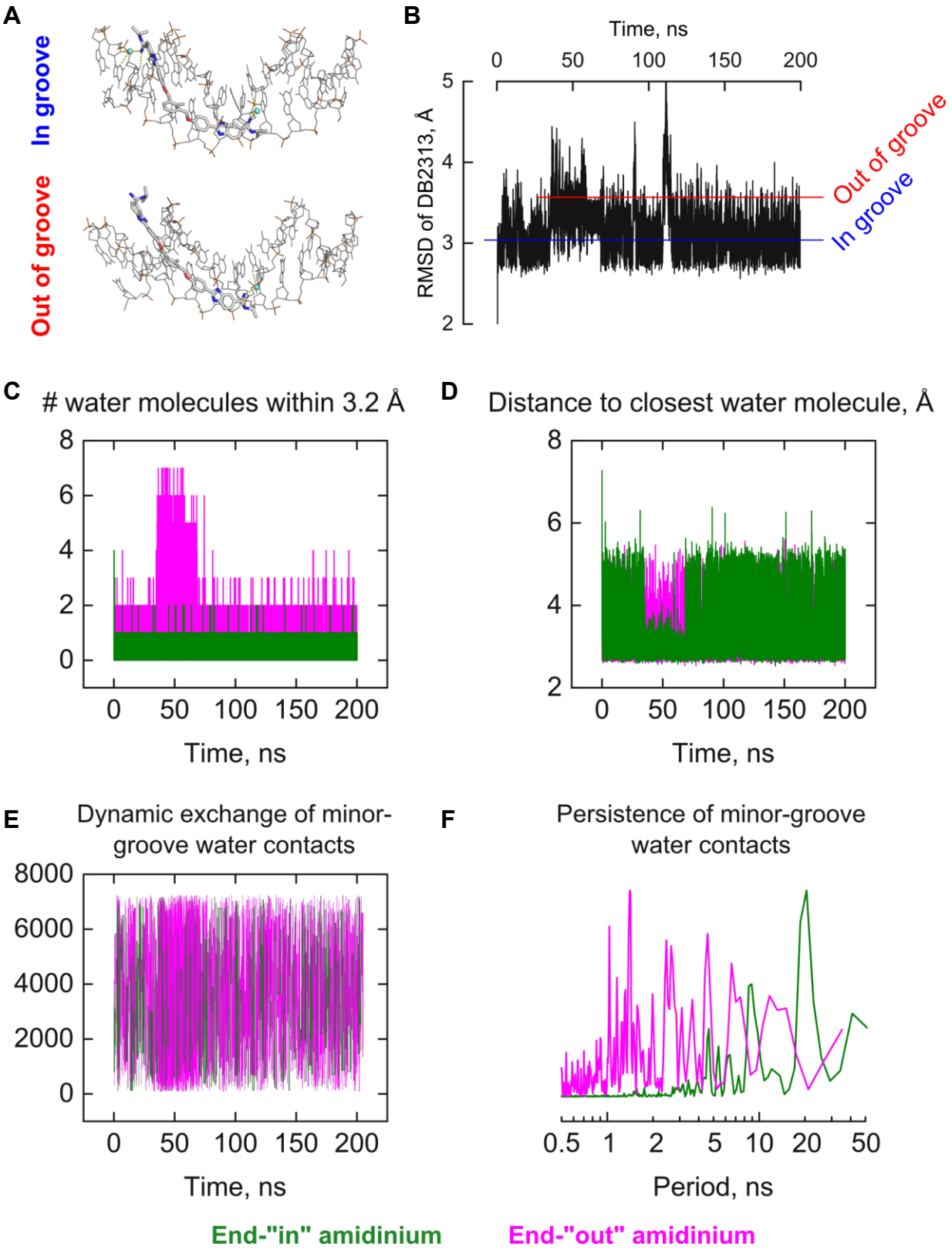
Supplementary Figure 1. PU.1 level and effects of PU.1 knockdown in cell lines



Supplementary Figure 1. PU.1 level and effects of PU.1 knockdown in cell lines.

(A) PU.1 protein expression after transduction with shPU.1_1, shPU.1_2, shPU.1_3 and shPU.1_4 in MOLM13 and BaF3 cells. **(B)** Quantitative RT-PCR analysis of *PU.1* expression in PU.1 URE$^{-/-}$ AML cells after transduction with shPU.1_1 (n=5), shPU.1_2 (n=5) or shPU.1_3 (n=2), normalized to *Gapdh* expression. Fold change compared to shCtrl is shown. **(C)** Quantitative RT-PCR analysis of *PU.1* expression in BaF3 and PU.1 URE$^{-/-}$ AML cells, normalized to *Gapdh* expression (n=4). **(D)** Cell proliferation assay of BaF3 cells after transduction with shPU.1_1, shPU.1_2 or shPU.1_3 (n=3). One representative experiment is shown. **(E)** Clonogenic capacity of BaF3 cells after transduction with shPU.1_1, shPU.1_2 or shPU.1_3 (n=3). Fold change compared to shCtrl is shown. **(F)** Apoptotic cell (Annexin-V^{+}DAPI^{-}) fraction in BaF3 cells after transduction with shPU.1_1, shPU.1_2 or shPU.1_3 (n=3). Fold change compared to shCtrl is shown. **(G)** PU.1 protein expression in THP1, MOLM13 and Kasumi-1 cells. **(H)** Quantification of PU.1 protein levels after transduction of MOLM13 cells with shPU.1_2 and/or pCAD PU.1 (n=4). Fold change compared to pCAD empty + shCtrl is shown. **(I)** Clonogenic capacity of MOLM13 cells after transduction with shPU.1_2 and/or pCAD PU.1 (n=4). Fold change compared to pCAD empty + shCtrl is shown. **(J)** Apoptotic cell (Annexin-V+DAPI-) fraction in MOLM13 cells after transduction with shPU.1_2 and/or pCAD PU.1 (n=4). Fold change compared to pCAD empty + shCtrl is shown. *P* values were determined using (B, H, J) one-way Anova or (C) 2-tailed Student's *t* test. * P<0.05

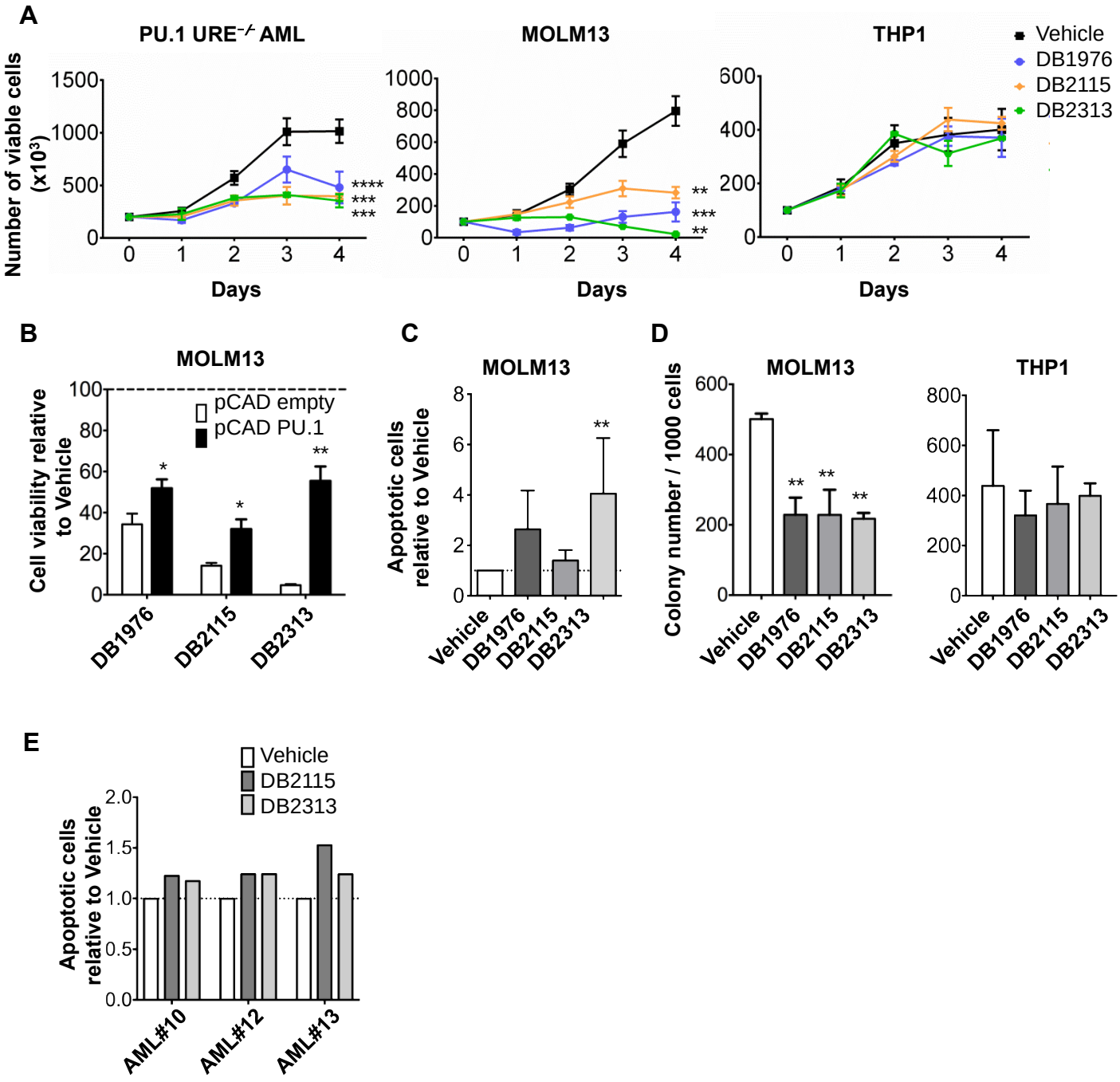
Supplementary Figure 2. Molecular dynamics simulations of minor groove recognition by DB2313



Supplementary Figure 2. Molecular dynamics simulations of minor groove recognition by DB2313

The docked structure of DB2313 with the λB motif (as shown in Figure 2B in the main text) was modeled by all-atom simulation for over 200 ns. To examine the role of hydration in DB2313/DNA binding, the AT binding site was positioned with one end at the DNA terminus. The intrinsic fraying of the DNA end would represent a pseudo-unbound state with respect to the compound. (**A**) Insertion of the amidiniums in the DNA minor groove is mediated by ordered water molecules (cyan spheres). However, the amidinium facing the solvent-accessible DNA terminus is significantly more mobile, alternately sampling an in-groove (end "in") and out-of-groove conformation with concomitant loss of the ordered water molecule (end "out"). **(B)** RMS deviation of the DB2313 from its initial configuration over 200 ns of molecular dynamics (MD) simulation. The blue and red lines denote the in-groove and out-of-groove configurations of the dynamic ensemble. **(C-F)** The hydration properties associated with the two cationic nitrogens at the terminal amidiniums of DB2313 in complex with the λB motif were further analyzed in terms of number of water molecules (C) within a nominal hydrogen-bonding range (3.2 Å) and distance to nearest water molecule (D) for the amidinium nitrogens. The end-"in" amidinium (magenta) refers to the moiety facing the terminus of the simulated DNA and the end-"out" counterpart (green) faces the center of the DNA duplex in Panel A. The "bump" and "pit" in Panels C and D correspond to the extended out-of-groove excursion sampled by the end-"out" amidinium. **(E)** Time-dependent, stochastic exchange of buried water-mediated contact (c.f. cyan spheres in Panel A). The water contact associated with the end-"in" amidinium exchanges at a substantially lower frequency than its end-"out" counterpart. **(F)** To quantify this difference, a discrete Fourier transform was performed on the final 100 ns of the data in Panels E and F, and presented as a function of period (reciprocal frequency). The end-"out" amidinium water exchange profile is clearly dominated by fast short-lived, ~10^{-9}-s exchanges that are absent for the end-"out" amidinium, indicating the persistence of the latter's hydration status. Taken together, these dynamics highlight the essentiality of molecular hydration in target recognition by the expanded heterocyclic diamidines, as previously observed with shorter diamidines.

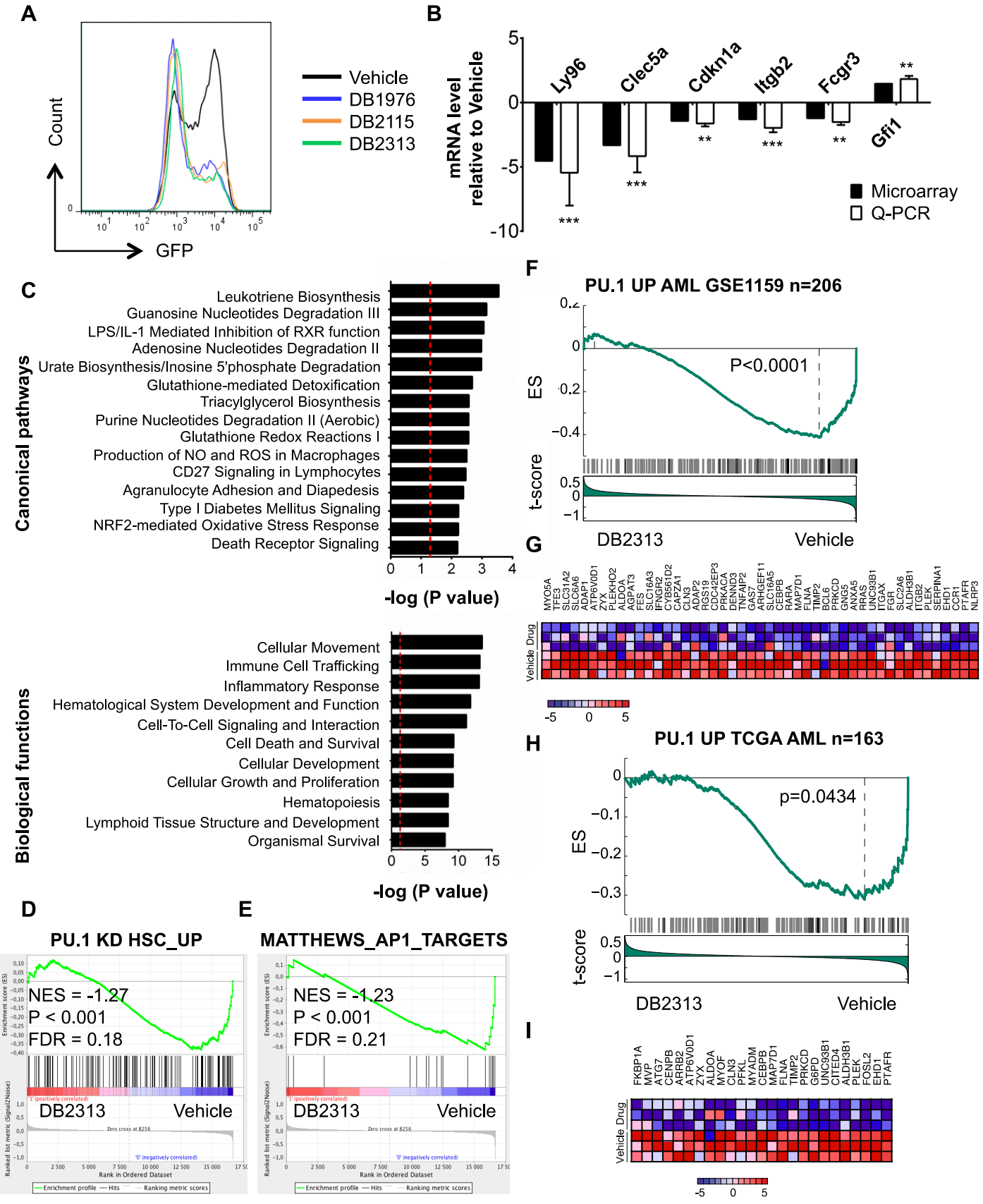
Supplementary Figure 3. Effects of PU.1 inhibition on leukemic cells



Supplementary Figure 3. Effects of PU.1 inhibition on leukemic cells

(A) Cell proliferation of PU.1 URE$^{-/-}$ AML, MOLM13 and THP1 cells after treatment with DB1976 (n=5), DB2115 (n=3), DB2313 (n=3). One representative experiment is shown. **(B)** Cell viability after PU.1 rescue in MOLM13 cells. MOLM13 cells transduced with pCAD PU.1 or pCAD empty as a control were plated in liquid culture with DB1976 (n=7), DB2115 (n=8), DB2313 (n=6) or Vehicle. Cell count was performed after 4 days. Fold change compared to Vehicle is shown. **(C)** Apoptotic cell (Annexin-V^{+}PI^{-}) fraction in MOLM13 cells after 72h of treatment with the small molecules (n=7). Fold change compared to Vehicle is shown. **(D)** Clonogenic capacities of MOLM13 cells (n=3) and THP1 cells (n=3) after treatment with the small molecules. **(E)** Apoptotic cell (Annexin-V+PI-) fraction in 3 primary AML samples after 72h of treatment with the small molecules. *P* values were determined using (A, C, D) one-way Anova or (B) 2-tailed Student's *t* test. * P<0.05, **P<0.01, ***P<0.001, ****P<0.0001

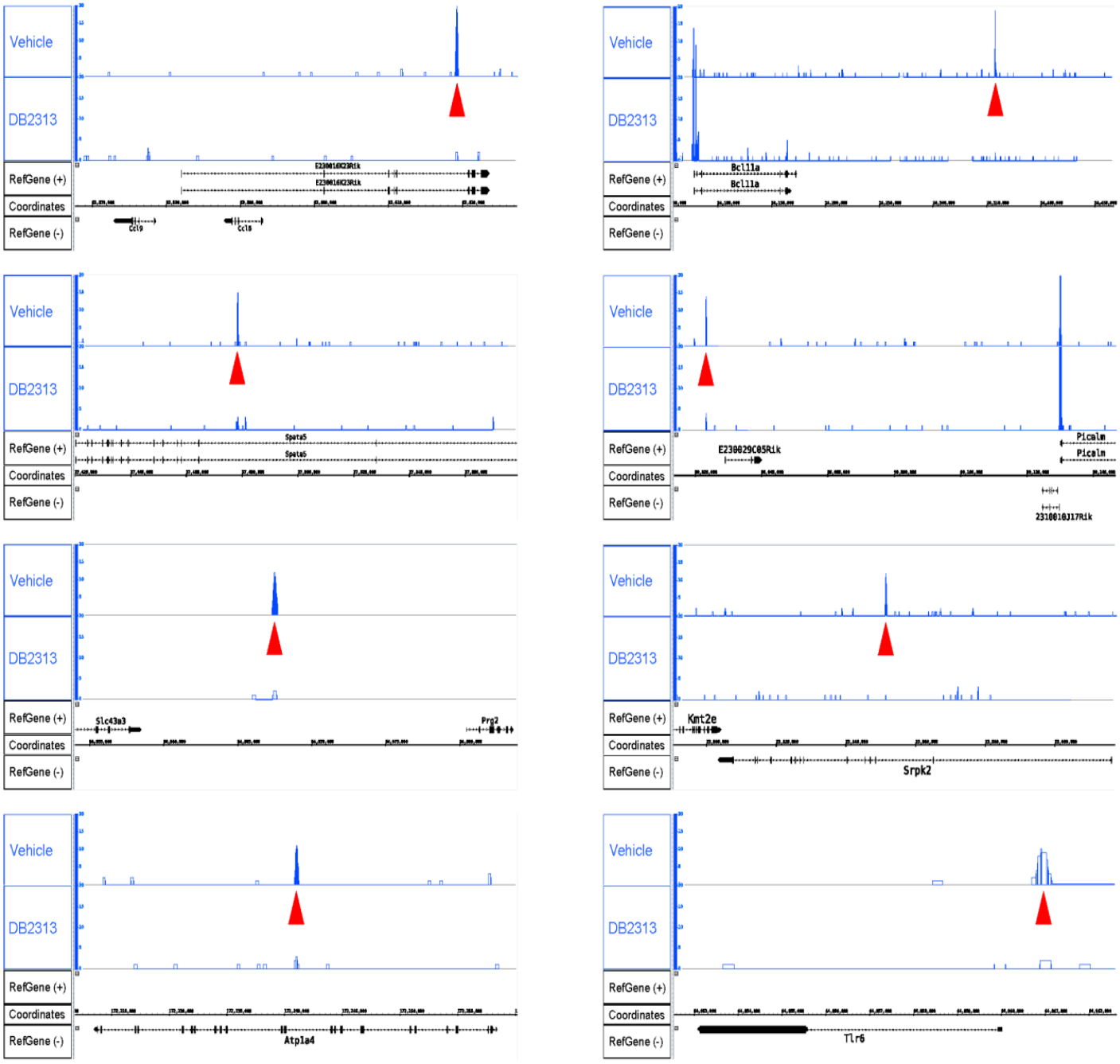
Supplementary Figure 4. Novel inhibitors show on-target PU.1 inhibitory activity



Supplementary Figure 4. Novel inhibitors show on-target PU.1 inhibitory activity.

(A) Histogram overlay showing the GFP intensity of bone marrow mononuclear cells isolated from PU.1/GFP mice after 4 days of treatment with PU.1 inhibitors. **(B-I)** Transcriptome analysis of PU.1 $URE^{-/-}$ AML cells after 24h of treatment with DB2313 versus vehicle (n=3). **(B)** Q-RT-PCR validation of several genes found deregulated by array. Expression was normalized to *Gapdh* (n=3). Fold change compared to Vehicle is shown. **(C)** Identification of canonical pathways and biological functions with significant enrichment after DB2313 treatment, by Ingenuity pathway analysis. The red dotted line represents the significance threshold (-log (P value) > 1.3). **(D, E)** Gene Set Enrichment analysis of genes expressed differently in DB2313 treated cells compared to the control. **(F)** GSEA enrichment plot of PU.1 positively regulated genes (regulon) against the global list of genes, from the GSE1159 AML network, ranked by the drug response (as measured by t-score of DB2313 vs. vehicle). **(G)** Heatmap of leading edge genes (subset of regulon contributing to the enrichment), showing row normalized relative expression. **(H)** GSEA enrichment plot of PU.1 positively regulated genes (regulon) against the global list of genes, from TCGA AML network, ranked by the drug response (as measured by t-score of DB2313 vs. vehicle). **(I)** Heatmap of leading edge genes (subset of regulon contributing to the enrichment), showing row normalized relative expression. *P* values were determined using (B) 2-tailed Student's *t* test or (D, E, F, H) according to ref (1). **P<0.01, ***P<0.001

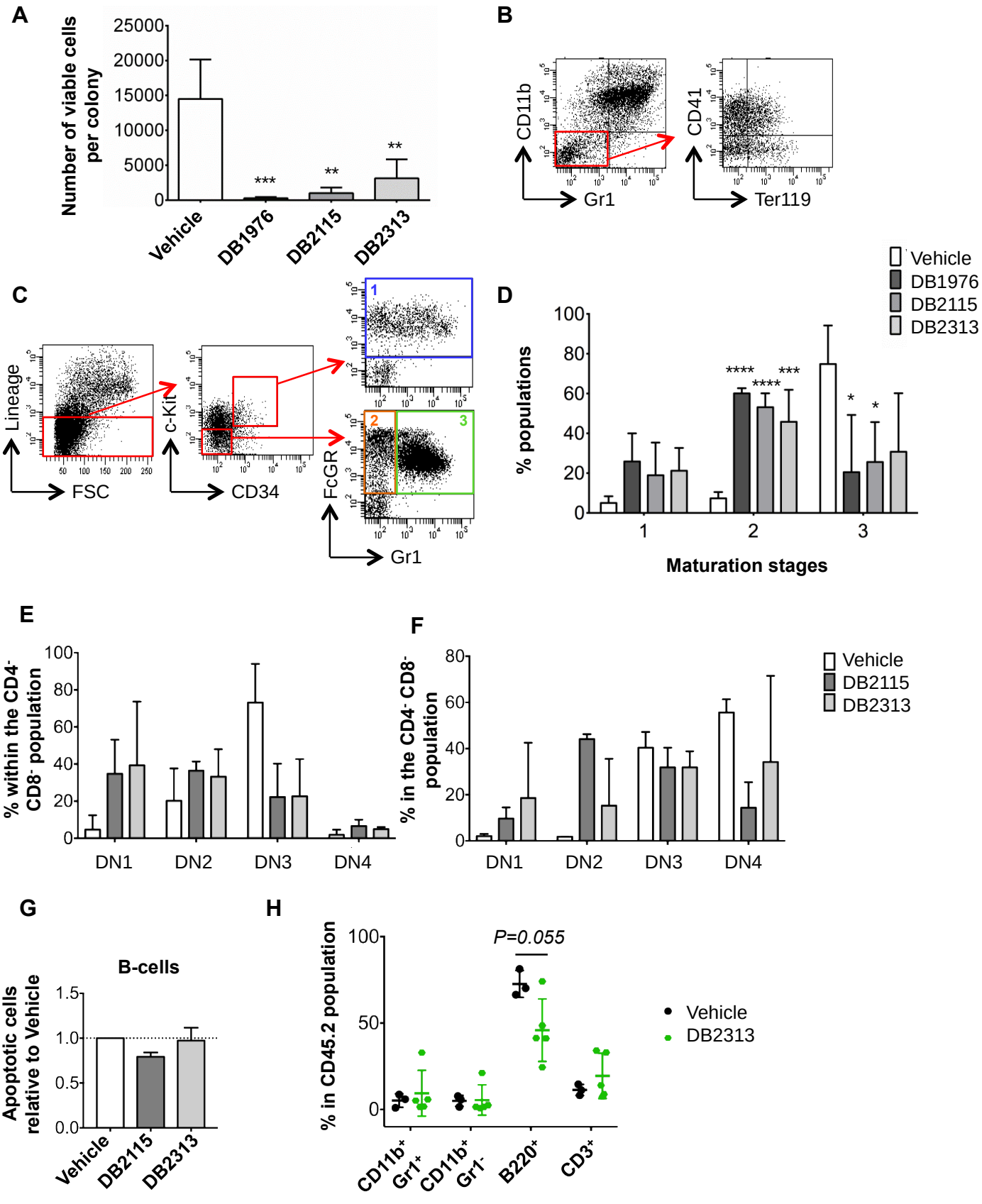
Supplementary Figure 5. Novel inhibitors decrease PU.1 binding to its target sites.



Supplementary Figure 5. Novel inhibitors decrease PU.1 binding to its target sites.

Examples of PU.1 peaks (indicated by red arrows) reduced or lost after DB2313 treatment (in comparison to vehicle-treated control) from PU.1 ChIP-seq analysis of PU.1 URE$^{-/-}$ AML cells after 24h of treatment with DB2313 or Vehicle. Please see Supplementary Table 2 for a complete list.

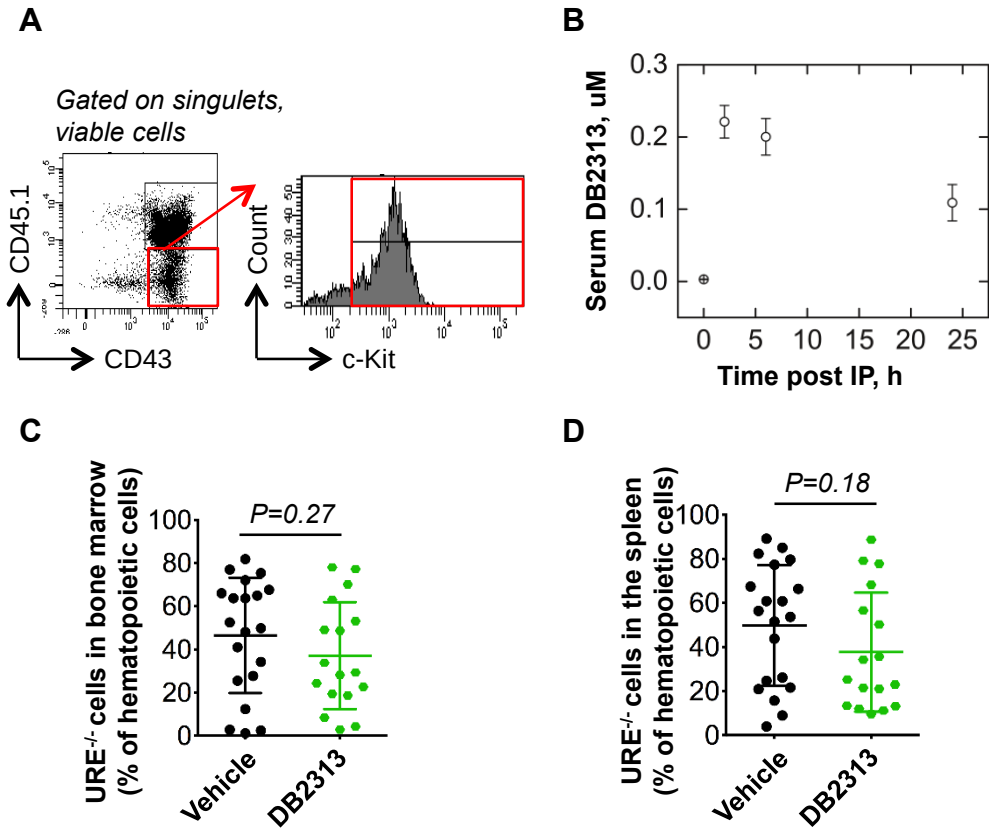
Supplementary Figure 6. Effects of PU.1 inhibition on normal hematopoiesis



Supplementary Figure 6. Effects of PU.1 inhibition on normal hematopoiesis.

(A) Number of viable cells per colony with PU.1 inhibitor treatment after plating LSK cells (n=4). **(B)** FACS plot showing the gating strategy used in Figure 6C, for assessing populations expressing Gr1, CD11b, Ter119 and CD41. **(C)** FACS plot showing the gating strategy to identify steps of granulocytic differentiation (with 1 = myeloblast-promyelocyte stages, 2 = myelocyte-metamyelocyte stages and 3 = mature granulocyte stage), based on CD34, c-Kit, Fcγr and Gr1 expression (2). **(D)** Percentage of populations in different maturation stages (1, 2 and 3) after treatment with PU.1 inhibitors (n=4). **(E-F)** Percentage of DN1 ($CD44^+CD25^-$), DN2 ($CD44^+CD25^+$), DN3 ($CD44^-CD25^+$), DN4 ($CD44^-CD25^-$) populations within the double negative $CD4^-CD8^-$ fraction after (E) 9 days (n=3) or (F) 18 days (n=2) culture of early thymic progenitors on OP9-Delta1 cells. **(G)** Apoptotic cell (Annexin-V+PI-) fraction in B-cells after 72h of treatment with the small molecules (n=3). Fold change compared to Vehicle is shown. **(H)** Blood analysis of each population ($CD11b^+Gr1^+$, $CD11b^+Gr1^-$, $B220^+$, $CD3^+$) in the CD45.2 fraction 20 Weeks after transplant of WT bone marrow cells treated with DB2313 (n=3 mice; vehicle group; n=5 mice, DB2313 group). Means ± SD are indicated by lines, each mouse is represented by an individual dot. *P* values were determined using (A, D) one-way Anova or (H) 2-tailed Student's *t* test. **P<0.01, ***P<0.001

Supplementary Figure 7. Treatment with PU.1 inhibitors leads to decreased tumor burden and increased survival *in vivo*.



Supplementary Figure 7. Treatment with PU.1 inhibitors leads to decreased tumor burden and increased survival in vivo.

(A) Gating strategy for assessing PU.1 URE$^{-/-}$ AML cell chimerism. **(B)** DB2313 measurement in the serum after intraperitoneal (IP) injection of wild type mice at different time points (n=2 mice). **(C-D)** PU.1 URE$^{-/-}$ AML cells transplant followed by IP treatment with either vehicle of DB2313. Chimerism of PU.1 URE$^{-/-}$ AML cells in the bone marrow (C) and in the spleen (D) 3 weeks after transplant (n=20 mice, vehicle group; n=17, DB2313 group; 3 independent experiments). Means ± SD are indicated by lines, each mouse is represented by an individual dot. *P* values were determined using 2-tailed Student's *t* test.

SUPPLEMENTARY TABLES

Supplementary Table 1. Characteristics of primary AML samples studied

AML#	AML type	Karyotypic/FISH anomalies	Mutations
AML#1	AML-M2	tri21	*FLT3*-ITD
AML#2	AML-M2	tri8	no
AML#3	AML-M4	No	*TET2, ASXL1*
AML#4	AML-M5	No	*DNMT3A, NPM1, NRAS*
AML#5	AML post MPN	ND	*JAK2, IDH2, TP53*
AML#6	Not specified	t(1;4)(p21;q21)	Unknown
AML#7	Not specified	No	*FLT3*-ITD
AML#8	AML-M5	No	*ASXL1, BCOR, RUNX1, SRSF2, TET2*
AML#9	AML-M5	del EGR1(5q31), del TP53(17p13.1)	*DNMT3A, TP53*
AML#10	Not specified	No	*DNMT3A*, *IDH1*, *MLL*-PTD, *FLT3*-TKD
AML#11	AML-M5	No	*DNMT3A*, *NPM1*, *FLT3*-ITD
AML#12	Not specified	Monosomy 7	*ASXL1, RUNX1, EZH2*
AML#13	AML post MDS	No	*ASXL1, EZH2, RUNX1, STAG2, TET2*

Supplementary Table 2. Complete list of PU.1 ChIP seq Peaks with PU.1 motif in Vehicle not called in the DB2313-treated sample.

(please see online supplemental table / Excel file)

Supplementary Table 3. Primer sequences

	Fwd (5'-3')	**Rev (3'-5')**
Q-RT-PCR		
Gapdh	CCAGCCTCGTCCCGTAGAC	GCCTTGACTGTGCCGTTGA
PU.1	AGAAGCTGATGGCTTGGAGC	GCGAATCTTTTTCTTGCTGCC
Csf1r	TGAAGATGCTAAAGTCCACGG	CCTTCGGAGAAAGTTGAGTAGG
Junb	CTGTGTCCCCCATCAACATG	TTCCGCTTCCGGCACTT
E2f1	GCCCTTGACTATCACTTTGGTCTC	CCTTCCCATTTTGGTCTGCTC
Ly96	CTTTTCGACGCTGCTTTCTC	ATCCATTGGTTCCCCTCAGT
Clec5a	CGAGCAGGAGCATACATTCA	CTACGATAAGCCCCGAGATG
Cdkn1a	GTACTTCCTCTGCCCTGCTG	TCTGCGCTTGGAGTGATAGA
Itgb2	TGCGGTGACAAAGAAGATGGTGAA	GTGCCCGGATGACAAAGGACTG
Fcgr3	TACACAGCACCAGTCCAAGC	ATGGATGGGGTGTCACTTGT
Gfi1	AGGAACGCAGCTTTGACTGT	CCTGTGTGGATGAAGGTGTG
ChIP		
Csf1r	TTCCCTTTCAGGCAACCTAA	CCCAGCTGCTAGTTCTGTGA
Junb	GCTGCCAGGCTTATTAGTCG	CATTGTGAGCCTAGGGGATG
E2f1	CAGGCTTTGGCACCAAATTCCCAA	GCAAGCCAGCAGACATCAGTTCAA
Myogenin	GAATCACATGTAATCCACTGGA	ACGCCAACTGCTGGGTGCCA

SUPPLEMENTARY METHODS

Synthesis of DB2313

All solvents and reagents were used without purification as acquired from commercial sources. Melting points were measured using a capillary melting point apparatus which are uncorrected. Progress of the chemical reactions were monitored by thin-layer chromatography on silica gel 60-F_{254} aluminum plates and detected under UV light. All NMR spectra were recorded employing a 400 MHz spectrometer, and chemical shifts (δ) are in ppm relative to TMS as internal standard. Electrospray ionization (ESI) Q-Tof and Orbitrap were used for the mass spectra measurements. Elemental analyses are within ±0.4 of the theoretical values. Compounds reported as salts frequently analyzed for fractional moles of water; the proton NMR showed the presence of the indicated solvent.

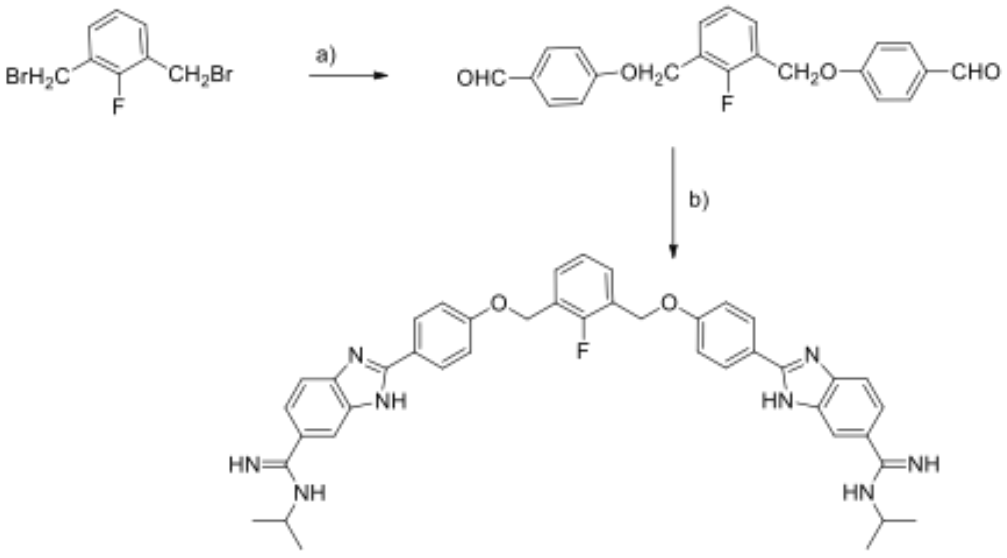
1, 3-Bis (4-formylphenoxymethyl) -2-fluorobenzene

A mixture of 1,3-bis(bromomethyl)-2-fluorobenzene (3) (1.41 g, 0.005 mole), 4-hydroxybenzaldehyde (1.22 g, 0.01 mole) and anhydrous. K_2CO_3 (2.07 g, 0.015 mole) in 10 ml DMF was heated at 45°C for 4 h [tlc (Hexane: EtOAc 8:2) monitored], diluted with ice water 70 ml, the precipitated white solid was filtered, washed with water, and dried. It was dissolved in DCM (75 ml), dried over anhydrous MgSO4, filtered, concentrated and triturated with cold hexane, filtered and dried under reduced pressure to yield a white solid 1.46 g (78%), mp 110-111°C; ^{1}H NMR (DMSO-d_6): 9.89 (s, 2H), 7.90 (d, 4H, J = 8.4 Hz),7.62 (t, 2H, J = 7.6 Hz),7.30 (t, 1H, J = 7.6 Hz), 7.26 (d, 4H, J = 8.4 Hz), 5.31 (s, 4H) ; ^{13}C NMR (DMSO-d_6): 191.8, 163.5, 159.5(J_{C-F} = 248.6 Hz), 132.3, 131.6 (J_{C-F} = 3.7 Hz) 130.5, 125.0 (J_{C-F} = 3.7 Hz), 123.9 (J_{C-F} = 14.7 Hz) 115.5, 64.4 (J_{C-F} = 4.0); MS: HRMS-ESIPOS.: Calcd. For: $C_{22}H_{17}FO_4$ Na *m/z* 387.1009(M^++Na), found *m/z* 387.1537; Anal. calcd. for: $C_{22}H_{17}FO_4$: C, 72.50; H, 4.70; Found: C,72.49; H,4.72.

3-Bis{4[4(5)-N-isopropylamidinobenzimidazolyl]phenoxymethyl]}-2-fluorobenzene tetrahydrochloride

A well stirred solution of 1, 3-bis (4-formylphenoxymethyl)-2-fluorobenzene (0.182 g, 0.0005 mole), 4-(*N*-isopropylamidino)-1, 2-phenylenediamine hydrochloride.0.2H_2O (4) (0.232 g, 0.001 mole) and 1, 4-benzoquinone (0.108 g, 0.001 mole) in anhydrous ethanol (40 ml) (under nitrogen) was heated at reflux for 8-10 h. The reaction mixture was cooled, concentrated to 10 ml and stirred in 50 ml acetone, filtered, washed with dry ether and dried to yield a hydrochloride salt. This salt was dissolved in a 1:1 mixture of hot ethanol-methanol (50 ml) and filtered, volume reduced to 20 ml and acidified with HCl-saturated ethanol (3 ml). After stirring overnight and diluting with anhydrous ether, filtered, washed with ether, and dried under reduced pressure at 70^0C (12 h) yielding purple-bluish grey solid 0.33 g (75%); mp > 320°C dec.; ^{1}H NMR (DMSO-d_6/65^0C): 9.68 (s, 1H), 9.67 (s, 1H), 9.53 (s, 2H), 9.12 (s, 2H), 8.41 (d, 4H, J = 8.4 Hz), 8.08 (s, 2H), 7.86 (d, 2H, J = 8.4 Hz), 7.67 (d, 2H, J = 8Hz), 7.65 (d, 2H, J = 8Hz), 7.37-7.32 (m, 5H), 5.34 (s, 4H), 4.10 (quintet, 2H, J = 6 Hz), 3.37-3.2 (vbs, benzimidazole NH), 1.31(d, 12H, J = 6 Hz); MS: HRMS-ESI-POS.: calc. for $C_{42}H_{42}FN_8O_2$ *m/z* 709.3415(M^++1), found *m/z* 709.4542; analysis calc. for $C_{42}H_{41}FN_8O_2$.4HCl.1.65H_20: C, 57.14; H, 5.52; N, 12.70; Found: C,57.35; H,5.67; N, 12.81.

Scheme 1.



Reagents and conditions: a) 4-hydroxybenzaldhyde, K_2CO_3, DMF, rt b) 3,4-diamino-*N*-(i-propyl)benzene-1-carboximidamide, benzoquinone, ethanol, reflux.

Patients characteristics

Characteristics of the primary AML samples studied are described in Supplementary Table 1. Conventional cytogenetics and FISH analysis were performed on patients cells, as well as detection of mutations in the following genes on genomic DNA: *ASXL1, CBL, DNMT3A, ETV6, EZH2, IDH1, IDH2, JAK2, KIT, MPL, NPM1, NRAS, PHF6, RUNX1, TET2, SETBP1, SF3B1, SRSF2, TP53, U2AF1, ZRSR2, MLL*-PTD according to standard methods. *FLT3*-ITD mutations were detected as previously described (5).

Molecular Dynamics Simulations

DB2313 was optimized at B3LYP/6-31*G using Spartan 10 (Wavefunction, Inc.). The ligand was assigned Gasteiger−Huckel charges using Autodock Tools 4.02 (6). The *d*-[5'-CCAAATAAAAGGAAG-3')(5'-CTTCCTTTTATTTGG-3')] λB duplex DNA sequence was generated with Tripos SYBYL-X1.2. The DNA sequence was docked with DB2313 using AutoDock Vina (6). One hundred runs were performed using the Lamarckian genetic algorithm (LGA) with no modifications of docking parameters, and the lowest-energy conformation selected for use. AmberTools16 was used to generate the requisite topologies for the simulation, and energy minimizations and molecular dynamics simulations were performed with Amber14. The forcefields used were as listed: GAFF parameters for the ligand (DB2313), and ff99bsc0/OL15 for DNA. The system was solvated with TIP3 waters as an octahedron with a minimum distance of 10.0 Å to the periodic simulation box. Counter ions were added to neutralize the net charge of the complex and to bring the overall salt concentration of the system to 0.15 M. The system was minimized (10,000 steps of steepest descent, 5,000 steps of conjugate gradient) with restraints applied to the complex. Afterwards, the system was minimized without restraints (15,000 steps of steepest descent, 10,000 steps of conjugate gradient). Particle mesh Ewald (PME) method was used to treat long-range electrostatic interactions; a cutoff for non-bonded interactions was set at 9 Å. SHAKE was used to constrain all bonds involving hydrogen atoms, with all bonds involving hydrogen atoms were constrained

using the LINCS algorithm. The system was gradually heated to 300 K over 50 ps in the NVT ensemble with the Langevin themostat. After heating to 300 K, the system was equilibrated (p = 1 atm, T = 300 K) using the NPT ensemble for 10.0 ns. Prior to production, 1-fs time steps were used. Afterwards, the system was run for 200 ns in an isothermal isobaric ensemble (T = 300K, p = 1 atm) with 2 fs time steps. Data was analyzed using the CPPTRAJ code in Amber16 and VMD.

Biosensor-SPR assays for compound binding affinity and PU.1-DNA complex inhibition by compound

SPR experiments were performed with a four-channel Biacore T200 optical biosensor system (GE Healthcare). A streptavidin-derivatized CM5 sensor chip was prepared for use by covalent linkage of streptavidin followed by conditioning with a series of 60 s injections of 1 M NaCl in 50 mM NaOH and extensive washing with HBS buffer [10 mM HEPES, 150 mM NaCl, 3 mM EDTA, and 0.05% P_{20} (pH 7.4)]. A 5'-biotin labeled hairpin λB DNA sequence (5'-CCAAATAAAAGGAAGTGAAACCAAGCTCTCTTGGTTTCACTTCCTTTTATTTGG-3' (hairpin loop underlined) was prepared in HBS buffer and immobilized on the flow cell surface by noncovalent capture as previously described (7). To determine the binding constant of the ligand with the λB DNA sequence, typically, a series of different compound concentrations (from 1 nM to 1 μM) were injected over the DNA sensor chip at a flow rate of 100 μL/min and the SPR response was followed for 3 min. This was followed by buffer flow to monitor the compound dissociation from the DNA complex and yielded a complete sensorgram for each compound concentration. After each cycle, the sensor chip surface was regenerated with a 10 mM glycine solution at pH 2.5 for 30 s followed by multiple buffer injections to yield a stable baseline for the following cycles. Kinetic analyses were performed by fitting the SPR sensorgram set by using a standard 1:1 kinetic model with integrated mass transport-limited binding parameters as described previously (8, 9).

For protein inhibition studies by the compounds, a 100 nM constant concentration of PU.1 protein was injected on the surface to essentially saturate the immobilized DNA binding sites. Graded concentrations

of compound were then added to the protein solution. The decrease in the protein binding signal in presence of different concentration of compounds was plotted against compound concentration as previously described to determine the inhibition by each compound (10).

DNA Footprinting

DNA occupancy by compounds and their conformational perturbations were probed in terms of protection of the DNA from DNase I as previously described (11). In brief, a DNA fragment harboring the λB motif was cloned into pUC19 and PCR-amplified using modified M13-based primers that were end-labeled with distinct fluorophores (Alexa Fluor 488 and HEX) at their 5' end. Following purification by agarose electrophoresis, the 206-bp fragments were saturated with 1 μM compound and subject to chemical or enzymatic modification. Digested samples were purified either by ethanol precipitation or using spin columns (Thermo Scientific), and analyzed by capillary electrophoresis (Macrogen, Seoul, South Korea). Peaks were indexed using a dimethyl sulfate-treated sample to mark purine nucleotides and a diagnostic DNase-hypersensitive peak general by a PU.1-bound (100 nM) control (12).

Lymphoid cell culture

For B-cell culture, wild type bone marrow cells were maintained in RPMI supplemented with 10% FBS, 50ng/ml rmIL-2, 10 ng/ml rmIL-4, 10 ng/ml rmIL-21, and 10ng/ml rmIL7 (all from Gemini), in presence of vehicle (water) or different compounds. Apoptosis was measured 72h after treatment.

For T-cell culture, early thymic progenitors sorted for Lin- (CD4, CD8, CD19, B220, CD11b, Gr1, Ter119), CD44+, CD25-, c-Kit+ expression were plated on OP9-Delta1 cells in α-MEM media (Gibco) supplemented with 20% FBS, 0.22% NaHCO3, 1% P/S, 5ng/ml rmFlt3-L, 2ng/ml rmIL-7 (Gemini), in presence of vehicle (water) or different compounds. Cells were analyzed by flow cytometry at day 9 and day18 for CD44, CD25, CD4, and CD8a expression.

Flow cytometry analysis and sorting

Apoptosis was measured with Annexin-V FLUOS staining kit (Roche), using Annexin-FITC and PI for cells treated with the compounds or Annexin-PE and DAPI for cells transduced with shRNAs. For other stainings, single cell suspensions were incubated with conjugated monoclonal antibodies for 20 minutes at 4°C, and washed in PBS 1X before acquisition on LSRII (Becton Dickinson) or sort on AriaII (Becton Dickinson) instruments. Analysis of FACS data was performed using the BD FACSDiva (Becton Dickinson) or FlowJo (Treestar) software. The following antibodies were used in the experiments: for mouse cells, c-Kit [2B8], Gr1 [RB6-8C5], CD11b [M1/70], CD41 [MWReg30], Ter119 [TER-119], CD4 [GK1.5], CD8a [53-6.7], CD19 [1D3], B220 [RA3-6B2], CD34 [RAM34], FcγR [93], Sca1 [D7], CD45.1 [A20], CD43 [R2/60], CD3 [145-2C11], CD44 [IM7], CD25 [PC61.5] (all from LifeTechnologies); for human cells, CD45 [HI30] (Becton Dickinson). Wild type LSK cells were sorted after exclusion of Lin^+ cells (expressing CD4, CD8a, CD19, B220, Ter119, CD11b, or Gr1).

Cytomorphology

Cell morphology was assessed after cytospin of 50-100 000 cells onto a glass slide (5 min at 500 rpm) and May-Grünwald Giemsa staining, according to standard protocols. Images were obtained using an EVOS FL Auto microscope (Life Technologies) with an objective at a 50x magnification.

Western blotting

1 to $2x10^6$ cells were lyzed in RIPA buffer, separated by electrophoresis on a 10% polyacrylamide gel and transferred to a nitrocellulose membrane (Maine manufacturing). Membranes were blocked for 1h using 5% milk in Tris-buffered saline containing 0.1% Tween-20 (TBS-T). Incubation with primary antibodies anti-PU.1 (1:1000, Santa-Cruz, SC-352) or anti-β tubulin (1:1000, Sigma-aldrich, T8328) was done overnight at +4°C. After 4 washes in TBS-T, membranes were incubated with secondary antibody goat anti-rabbit (1:5000, Santa-Cruz, SC-2004) or anti-mouse (1:5000, Santa-Cruz, SC-2005) conjugated to

horseradish peroxidase for 1h at RT. Bands were visualized using with the ECL system (Thermoscientific).

Molecular Modeling and Docking

DB2313 was optimized at the B3LYP/6-31*G level of theory using Spartan 10 software (Wavefunction, Inc.). The minimized ligand was assigned Gasteiger−Huckel charges by using Autodock vina 4.02 (6). The *d*-[5'-CCAAATAAAAGGAAG-3')(5'-CTTCCTTTTATTTGG-3')]λB duplex DNA sequence was generated from the biopolymer-build DNA double helix module from the Tripos SYBYL-X1.2 software package (13). Duplex DNA sequence was docked with minimized structure of DB2313 using AutoDock vina 4.02 (6). The center of the macromolecule is the grid center with a grid size of 68 Å×70 Å×120 Å and a grid spacing of 0.375 Å. Docking runs were performed using the Lamarckian genetic algorithm (LGA) with no modifications of docking parameters. LGA was used because of the existence of rotatable bonds in the ligands and to evaluate the correct conjugate DNA conformation, as it is known to reproduce various experimental ligand−DNA complex structures. Initially, we used a population of random individuals (population size of 150), a maximal number of 2500000 energy evaluations, a maximal number of evaluations of 2700, and a mutation rate of 0.02 fs. Two hundred independent flexible docking runs were conducted for each ligand, and then the lowest-energy dock conformation obtained from the Flexible docking was resubmitted for rigid docking to remove the internal energy of the ligand (steric clashes) and retain the hydrogen bonding interaction with ds-DNA bases.

Interactome analyses (ARACNe Networks):

To generate AML transcriptional networks we processed microarray gene expression profiles independently from 3 independent series AML primary samples including TCGA (183 AML samples, Affymetrix U133 Plus2), GSE13159 (MILE study, 542 AML samples, Affymetrix U133 Plus2),

GSE1159 (293 AML samples, Affymetrix U133A) (14-18). Expression profiles were normalized using GC-RMA and processed through the ARACNe algorithm as described in Margolin et al (19-21).

The transcriptional activity of PU.1 was inferred by interrogating the AML networks with the DB2313-induced signature. Specifically the PU.1 "regulon" as predicted by the ARACNe algorithm was partitioned in positive and negative based on positive or negative spearman correlation ($P<0.05$) between expression levels of PU.1 and the target gene across all samples of the dataset used to generate the network. PU.1 positive and negative regulons were tested for enrichment in the drug induced expression signature by Gene Set Enrichment Analysis (GSEA) (1).

ChIP-seq analysis

$10x10^6$ PU.1 $URE^{+/-}$ AML cells were treated with either Vehicle or 660nM of DB2313 for 24h. Following the ChIP experimental procedure described in the main text, immunoprecipitated chromatin was submitted to the Albert Einstein College of Medicine Epigenomics Shared Facility for library preparation and sequencing on an Illumina NextSeq500 instrument. Single-ended 150bp Illumina [NextSeq500] reads were aligned to the mouse genome (mm10) using bowtie2 (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml, version version 2.3.1) with default parameters after trimming adapters and low quality reads using bbduk from the BBTools suite (http://jgi.doe.gov/data-and-tools/bbtools/, version 36.19) (using the parameters minlen=31 qtrim=rl trimq=10 ktrim=r k=25 mink=11). Adjustment of read numbers in IP samples and background samples was performed independently by random resampling. ChIP peaks were called with macs2 (https://github.com/taoliu/MACS, version 2.1.1.20160309) using default parameters and the respective input controls as background. Motif enrichment analysis was performed using homer (http://homer.ucsd.edu) with default parameters. Identification of the top homer motif within peaks was performed using the script annotatePeaks.pl distributed with homer. To identify bona fide PU.1 peaks that diminish with DB2313 treatment, we used the by far most highly enriched de novo motif identified by homer in the vehicle control peaks compared with the input control (p=1e-844, present in 41.1% of

targets, but only in 4.3% of background sequences) - which contains a PU.1 consensus motif - to identify peaks with this motif in DB2313- and vehicle-treated PU.1 ChIP-seq samples.

Bone marrow transplantation

Total BMC from C57/BL6 WT CD45.2$^+$ mice were isolated from the tibiae, femurs and pelvic bones, treated with ACK buffer pH 7.4 to lyse red blood cells, and cultured in presence of vehicle (water) or different compounds. Viable cells were counted 2 days later and transplanted retroorbitally into lethally irradiated (925 rad) C57/BL6 WT CD45.1$^+$ mice (2.106 cells transplanted per mouse). Analysis of Gr1, CD11b, B220 and CD3 expression in CD45.2$^+$ peripheral blood cells was performed by flow cytometry 20 weeks after transplant.

Determination of serum compound concentrations

To each 30 µL of serum sample, 4 µL of a 1 mM an internal standard (DB2373) was added, followed by 200 µL of extraction solution (7:1 MeOH-H_2O containing 0.1% TFA solution). The mixture was extracted by rigorous vortexing for 30 s and then centrifuged at 2800 × g for 1 min. The supernatant was transferred to a new tube and dried under vacuum without heat. Standards containing analytically weighed DB2313 were spiked with DB2373 and treated similarly.

API-MS detection was achieved using an AB SCIEX API 3200^{TM} LC/MS/MS triple quadrupole mass spectrometer equipped with an orthogonal Turbo electrospray ion source and Agilent 1200 Series HPLC System. The samples were delivered to the ionization source by the autosampler of Agilent 1200 HPLC system. A C18 column (3 µm particle size, 3×100 mm, Phenomenex, Torrance CA, USA) was used for the chromatographic separation. The injection volume was 5 µL. The HPLC eluents were H_2O with 0.1% formic acid (A) and methanol with 0.1% formic acid (B). The initial gradient condition was 90% A and 10% B, holding on 10% B for 1 min, changing linearly to 60% B in 4 min, standing on 60% B for 2 min, then changing linearly to 90% B in 2 min, standing on 90% B for 2 min, decreasing linearly to 10% B within 1 min and equilibrating at 10% B for 5 min, giving a total analysis time of 17 min. After

analysis the column was washed with the above procedure, giving background chromatographic spectra. The eluent flow rate was 0.3 mL/min. Data were processed using the Analyst v1.5 software from AB SCIEX.

All compounds were detected using positive ionization in multiple reaction monitoring (MRM) mode (ion pair 355/294 for DB2313 and 346/294 for DB 2373). The analyte-dependent and ESI source parameters were selected to optimize detection of DB2313 (ion pair 355/294). The experimental parameters were set as follows: ion spray voltage 4500 V, ion source gas GS1 45 psi, ion source gas 2 (GS2) 35 psi, ion source temperature was 450 °C, with the interface heater turned ON. The declustering potential (DP) was optimized at 40 V and entrance potential (EP) was 6 V. Collision energy (CE) was 30V and collision gas was 6 psi. Deflector was preset to 200 V and CEM was operated at 2400 V. The dwelltime was 200ms. The system was operated under control of Analyst 1.5 software.

SUPPLEMENTARY REFERENCES

1. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America.* 2005;102(43):15545-50.
2. Guibal FC, Alberich-Jorda M, Hirai H, Ebralidze A, Levantini E, Di Ruscio A, Zhang P, Santana-Lemos BA, Neuberg D, Wagers AJ, et al. Identification of a myeloid committed progenitor as the cancer-initiating cell in acute promyelocytic leukemia. *Blood.* 2009;114(27):5415-25.
3. Casitas A, Canta M, Sola M, Costas M, and Ribas X. Nucleophilic aryl fluorination and aryl halide exchange mediated by a Cu(I)/Cu(III) catalytic cycle. *Journal of the American Chemical Society.* 2011;133(48):19386-92.
4. Ates-Alagoz Z, Alp M, Kus C, Yildiz S, Buyukbingol E, and Goker H. Synthesis and potent antimicrobial activities of some novel retinoidal monocationic benzimidazoles. *Archiv der Pharmazie*. 2006;339(2):74-80.
5. Kiyoi H, Naoe T, Nakano Y, Yokota S, Minami S, Miyawaki S, Asou N, Kuriyama K, Jinnai I, Shimazaki C, et al. Prognostic implication of FLT3 and N-RAS gene mutations in acute myeloid leukemia. *Blood.* 1999;93(9):3074-80.
6. Trott O, and Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry.* 2010;31(2):455-61.
7. Nguyen B, Tanious FA, and Wilson WD. Biosensor-surface plasmon resonance: quantitative analysis of small molecule-nucleic acid interactions. *Methods (San Diego, Calif).* 2007;42(2):150-61.

8. Liu Y, Chai Y, Kumar A, Tidwell RR, Boykin DW, and Wilson WD. Designed compounds for recognition of 10 base pairs of DNA with two at binding sites. *Journal of the American Chemical Society.* 2012;134(11):5290-9.
9. Paul A, Chai Y, Boykin DW, and Wilson WD. Understanding mixed sequence DNA recognition by novel designed compounds: the kinetic and thermodynamic behavior of azabenzimidazole diamidines. *Biochemistry.* 2015;54(2):577-87.
10. Munde M, Poon GM, and Wilson WD. Probing the electrostatics and pharmacological modulation of sequence-specific binding by the DNA-binding domain of the ETS family transcription factor PU.1: a binding affinity and kinetics investigation. *Journal of molecular biology.* 2013;425(10):1655-69.
11. He G, Tolic A, Bashkin JK, and Poon GM. Heterogeneous dynamics in DNA site discrimination by the structurally homologous DNA-binding domains of ETS-family transcription factors. *Nucleic acids research.* 2015;43(8):4322-31.
12. Poon GM. Sequence discrimination by DNA-binding domain of ETS family transcription factor PU.1 is linked to specific hydration of protein-DNA interface. *The Journal of biological chemistry.* 2012;287(22):18297-307.
13. Sanner MF. Python: a programming language for software integration and development. *Journal of molecular graphics & modelling.* 1999;17(1):57-61.
14. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *The New England journal of medicine.* 2013;368(22):2059-74.
15. Haferlach T, Kohlmann A, Wieczorek L, Basso G, Kronnie GT, Bene MC, De Vos J, Hernandez JM, Hofmann WK, Mills KI, et al. Clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: report from the International Microarray Innovations in Leukemia Study Group. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2010;28(15):2529-37.
16. Kohlmann A, Kipps TJ, Rassenti LZ, Downing JR, Shurtleff SA, Mills KI, Gilkes AF, Hofmann WK, Basso G, Dell'orto MC, et al. An international standardization programme towards the application of gene expression profiling in routine leukaemia diagnostics: the Microarray Innovations in LEukemia study prephase. *British journal of haematology.* 2008;142(5):802-7.
17. Stirewalt DL, Meshinchi S, Kopecky KJ, Fan W, Pogosova-Agadjanyan EL, Engel JH, Cronk MR, Dorcy KS, McQuary AR, Hockenbery D, et al. Identification of genes with abnormal expression changes in acute myeloid leukemia. *Genes, chromosomes & cancer.* 2008;47(1):8-20.
18. Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, Beverloo HB, Moorhouse MJ, van der Spek PJ, Lowenberg B, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. *The New England journal of medicine.* 2004;350(16):1617-28.
19. Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R, and Califano A. Reverse engineering of regulatory networks in human B cells. *Nature genetics.* 2005;37(4):382-90.
20. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Dalla Favera R, and Califano A. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC bioinformatics.* 2006;7 Suppl 1(S7.
21. Margolin AA, Wang K, Lim WK, Kustagi M, Nemenman I, and Califano A. Reverse engineering cellular networks. *Nature protocols.* 2006;1(2):662-71.