Brief Communication

Hematopoietic Stem Cell Transcription Factor PU.1 with Mutated β3/β4 Domain Selectively Elicits Myeloid Differentiation

Pallavi Gupta[1], Gurudutta U Gangenahalli[1], Daman Saluja[2], Yogesh K Verma[1], Jyoti Zack[2] Vimal K Singh[1], Neeraj K Satija[1] & Rajendra P Tripathi

[1]Stem Cell & Gene Therapy Research Group, Institute of Nuclear Medicine & Allied Sciences (INMAS), DRDO, Delhi-110054, India. [2]Ambedkar Center for Biomedical Research, University of Delhi, Delhi-110007, India.

Corresponding authors:

Gurudutta U Gangenahalli
Chief, Stem Cell & Gene Therapy Research Laboratory
Institute of Nuclear Medicine and Allied Sciences (INMAS)
DRDO, Delhi-110054, INDIA
Tel: 91-11-23905144
Fax: 91-11-23919509
Email: gugdutta@rediffmail.com

Daman Saluja
Ambedkar Centre for Biomedical Research (ACBR)
University of Delhi, Delhi-110007, INDIA
Tel: 91-11-27666272
Fax: 27666248
Email: dsalujach@yahoo.com

In many hematopoietic malignancies such as lymphomas and leukemias, aberrant differentiation is the major feature of the malignant phenotype that often results from a single genetic alteration and provides a site-specific target for therapy. Therefore, targeting of protein-protein interactions that have been identified as mediators of transcriptional repression that blocks normal hematopoietic differentiation holds great promise for therapeutic applications. An example is GATA-1, a critical erythroid transcription factor that is capable of suppressing the myeloid phenotype by down-regulating PU.1 in a dose-dependent manner. PU.1 is the most crucial transcription factor known to be required for myeloid differentiation of hematopoietic stem/progenitor cells. Its reduced expression correlates with a bad prognosis and immature phentoype in acute myeloid leukemia (AML). GATA-1 downregulates PU.1 by interacting though its C-terminal zinc finger region with the β3/β4 region of PU.1 and displacing its coactivator c-Jun. We hypothesize that disruption of PU.1-GATA-1 interaction by mutating the β3/β4 region of PU.1 may prevent its GATA-1-mediated repression, which in turn will upregulate PU.1 expression and hence myelopoiesis. Our analysis of the PU.1 mutants, Lys240Arg and Tyr244Ala, revealed that they exhibit an increase in myelopoiesis *in vitro*. Thus our data have implications for the prospect of targeting PU.1-GATA-1 interaction for therapeutic intervention.

Transcription factor PU.1 (Spi-1), which is encoded by the gene *Sfpi1* (spleen focus-forming proviral integration 1), belongs to ETS family of transcription factors. It is physiologically expressed in all hematopoietic cell lineages except T-cell lineage (**reviewed in Gangenahalli et al., 2005**). It recognizes a purine-rich sequence motif around a minimal core consensus, 5'-G/AGAA-3', in transcriptional promoters of various genes expressed primarily in myeloid and B-lymphoid lineages. PU.1 has been shown to regulate the expression of a growing list of genes in each of these lineages. It can activate the expression of myeloid cytokine receptors such as GM-CSFR, M-CSFR, GM-SCFR, mannose receptor and scavanger receptor and scavanger receptor, as well as lymphoid-specific interleukin-7 receptor. It can activate the expression of myeloid cytokine receptors such as granulocyte/macrophage colony stimulating factor α (GM-CSFRα), macrophage CSF receptor (M-CSFR), granulocyte CSF receptor (G-CSFR), mannose receptor and scavenger receptor, as well as the lymphoid-specific interleukin-7 receptor α (IL-7Rα)**.** PU.1 binding sites also exist in the enhancer of interleukin-4 (IL-4) gene in mast cells and in the enhancers/promoters of genes expressed during maturation and differentiation of B-lymphoid cells such as genes encoding the heavy, joining, and light chains of immunoglobulins. Disruption of its function blocks myelopoiesis both *in vitro* and *in vivo*. The PU.1 knockout mice show complete absence of mature monocytes and neutrophils, whereas graded reduction in the level of its expression can induce acute myeloid leukemia (AML) **(Rosenbauer et al., 2004)**. However, the retroviral restoration of PU.1 expression has been shown to rescue myeloid differentiation of mutant progenitors and AML blasts.

The PU.1 protein contains three distinct functional domains (**Fig. 1a**). The amino terminus contains both a highly acidic and a glutamine-rich transactivation domain (**Klemsz, 1996**). The carboxy-terminal region encompasses the evolutionarily conserved DNA-binding domain (also known as ETS domain) that contains approx. 85 amino acids (**Klemsz, 1990**). The transactivation and ETS domains are linked by a PEST (P-proline, E-glutamic acid, S-serine, T-threonine) domain also referred to as the potential protein degradation domain (**Rogers, 1986**). It has been demonstrated that transcription factor GATA-1 suppresses the myeloid gene expression program by inhibiting the PU.1 activity in a dose-dependent manner **(Nerlov et al., 2000)**. This inhibition takes place though a direct protein-protein interaction wherein the C-terminal zinc finger of GATA-1 physically binds with the β3/β4 region in the DNA binding domain of PU.1 and displaces its coactivator c-Jun (**Fig. 1b**). c-Jun acts as an important coactivator of PU.1 during the gene regulation of various myeloid promoters such as M-CSF receptor and macrosialin (**Behre,**

1999). Thus, GATA-1-mediated repression of PU.1 leads to a block in myeloid differentiation and could contribute to the malignant phenotype in some types of immature acute myeloid leukemia. Therefore, in order to prevent GATA-1-mediated repression of PU.1, we considered mutating the β3/β4 region of PU.1 so as to abolish its interaction with the repressor protein GATA-1. Here we have used a combined approach involving a detailed sequence and structural analysis of PU.1 protein followed by mutational analysis to identify PU.1 residues (in the β3/β4 region) critical for its interaction with GATA-1.

The molecular modeling and *in silico* docking techniques were used to study the specific molecular contacts between PU.1 and GATA-1 and to predict PU.1 mutants with reduced GATA-1 binding ability. The three-dimensional (3D) models of both PU.1 and GATA-1 were generated though a fully automated comparative protein homology modeling server SWISS-MODEL **(Schwede et al., 2003; Arnold et al., 2006)** (**Fig. 2a,b**) and subsequently verified for their accuracy in stereochemistry, non-bonded atomic interactions, 3D profile and protein volume using WHAT IF program **(Vriend, 1990; Hooft et al., 1996)** (**Table 1**). Prior to docking, the active site residues were predicted in the C-terminal Zn finger of GATA-1 by using ProMateus **(Neuvirth et al., 2007)**, a general protein binding site analysis web tool and the solvent accessibility surface areas of the residues in interacting regions of both the proteins were calculated using a web-based tool Parameter Optimized Surfaces (POPS) in order to filter key interactions between molecules due to the burial of surface area upon complex formation (**Tables 2 and 3**) **(Cavallo et al., 2002)**. The protein-protein docking was performed to generate minimum-energy complexes of PU.1 and GATA-1 using Hex 4.5 software (**Fig. 2c).** This led us to identify amino acid residues in PU.1 that were thought to be critical for its physical interaction with GATA-1, *viz.* Lys237, Val238, Lys239 and Lys240.

Further, in order to identify the PU.1 residues specifically involved in this interaction out of the predicted ones, the sequence analysis was performed. A PIR BLAST search followed by CLUSTAL W (version 1.81) multiple alignment of PU.1 homologues revealed that the β3/β4 region in their DNA binding domain was highly conserved **(Altschul et al., 1989; Higgins et al., 1996)**. However, certain residues were not so strictly conserved and such residues might be responsible for different functions in different proteins. For instance, in human Spi-B protein, Arg substituted the highly conserved Lys240. Spi-B is also an ETS family transcription factor that is most closely related to PU.1 in its DNA binding domain (Rey et al., 19). Rekhtman et al have reported that Spi-B also interacts with

GATA-1 but the interaction between them is much weaker than that between PU.1 and GATA-1. The weaker interaction between Spi-B and GATA-1 could be attributed to Arg240, which because of its bulky guanidinium group might cause steric hindrance to GATA-1. Hence, we predicted a mutation Lys240Arg in the β3/β4 region of PU.1 that would likely have reduced GATA-1 binding ability. We also decided to substitute Lys240 residue with Ala in order to determine whether a positively charged residue at this position is necessary for interaction with GATA-1. Since Ala is a small-sized uncharged amino acid residue, it would not affect the overall structural conformation of a protein.

Furthermore, only Tyr244 residue in the β3/β4 region of PU.1 was found to be significantly involved in intramolecular hydrogen bonding. In order to determine the role of Tyr244 in PU.1-GATA-1 interaction, we again performed protein-protein docking between GATA-1 and PU.1 mutant in which Tyr244 was substituted Ala (**Fig. 2d**). We observed a significant change in intramolecular H-bond lengths in PU.1 mutant Tyr244Ala (**Table 4**). Such changes in the intramolecular H-bonding of PU.1 are likely to induce a change in the conformation of its β3/β4 region, which probably would disrupt its interaction with GATA-1. Therefore, a third mutation Tyr244Ala was predicted. In order to validate these predictions, the above-mentioned mutants were generated by PCR-based site-directed mutagenesis and evaluated for enhanced myelopoiesis. The precise regulation of wild type PU.1 and its mutants in host cells is crucial as their deregulated overexpression would certainly lead to oncogenic conditions, thus necessitating the use of regulatable expression system. Therefore, prior to mutagenesis, wild type PU.1 was subcloned in an inducible mammalian gene expression vector, pTRE2 that uses a doxycycline-dependent transcriptional switch for controlling gene expression **(Gossen et al., 1994; 1995)**.

A T-lymphoblast cell line Jurkat E6.1 that lacks endogenous expression of PU.1 was used to study the effects of regulated overexpression of PU.1 mutants. The wild type PU.1 as well as mutants were transfected and stable cell lines were made. Since the point mutations in PU.1 were made in its DNA binding domain (β3/β4 region), we determined the DNA binding ability of these mutants (Lys240Arg, Lys240Ala and Tyr244Ala) through electrophoretic mobility shift assay (EMSA). Double stranded oligonucleotides containing PU.1 consensus site in the M-CSF receptor promoter were used as probes, which were end-labeled with [γ-^{32}P] dATP. Different concentrations of nuclear extract of cells expressing wild type PU.1 as well as mutants were incubated with the

labeled probe. A shift was observed with all concentrations of PU.1, however, the protein: DNA binding was found to increase with increasing concentrations of the protein (**Fig. 3a,c**). Thus DNA binding ability of PU.1 was not disrupted upon making mutations in its DNA binding domain. In order to determine the specificity of PU.1 protein for its DNA binding site, a competitive EMSA was performed using excess concentrations of unlabeled specific and non-specific probes (**Fig. 3b**). On using 1:2 and 1:10 concentrations of labeled:unlabeled specific probe, no shift was observed because the excess of unlabeled specific probe competed with labeled specific probe. However, in case of non-specific probe, a clear shift was observed at both 1:2 and 1:10 concentrations, as the non-specific probe did not compete with labeled specific probe. These results suggest that PU.1 binds to its specific DNA binding site.

The effects of overexpression of wild type PU.1 and mutants in Jurkat E6.1 cells were determined by flow cytometry for myeloid-cell specific markers CD33 and CD116 (GM-CSF receptor α). An increase was observed in CD33 and CD116 expression in cells overexpressing wild type PU.1 and mutants after 72 h of induction, but there was no change in the expression of T-cell marker CD3. The maximum increase in CD33 expression was found in case of mutants Tyr244Ala (17% more than wild type PU.1) ($P=0.05$) and Lys240Arg (13%) ($P=0.02$) as compared to the wild type PU.1 (**Fig. 4a,b**). Similarly, we found an increase in CD116 expression in cells overexpressing mutants Lys240Arg and Tyr244Ala by 45% ($P<0.01$) and 42% ($P<0.01$) respectively as compared to wild type PU.1 (**Fig. 4c,d**). This suggests that the mutants Tyr244Ala and Lys240Arg have a higher potential of generating myeloid cells and this could be due to their weaker interaction with the interacting protein GATA-1. Interestingly, the overexpression of wild type PU.1 and its mutants had no effect on the expression of T-lymphocyte marker CD3, which is characteristic of Jurkat cells (**Fig. 5**). This is in accordance with the previously established fact that expression of PU.1 gene is not required for the development of T-cells **(Spain et al., 1999)**.

We also assessed the effects of overexpression of wild type PU.1 and its mutants in $CD34^+$ human hematopoietic stem cells (HSCs) isolated from bone marrow by determining their ability to express myeloid marker CD33 after 96 h of doxycycline induction. A slight increase was observed in the expression of CD33 in cells overexpressing wild type PU.1 and mutants Lys240Arg and Lys240Ala, but there was a significant (32%) ($P<0.01$) increase in CD33 expression in case of HSCs overexpressing mutant Tyr244Ala (**Fig. 4e,f**). This suggests that the mutant Tyr244Ala has the highest potential for generating myeloid cells. This result is also in correlation with our *in silico* studies. The

structural changes inflicted on PU.1 by the mutation Tyr244Ala might be responsible for its reduced interaction with GATA-1 protein, which in turn might have led to an increase in the expression of myeloid marker CD33 in the HSCs expressing Tyr244Ala mutant.

Thus, we conclude that overexpression of PU.1 and its mutants Lys240Arg and Tyr244Ala increases myeloid lineage formation in human $CD34^+$ progenitor cells as well as cells that are already committed to other lineages. Therefore, we propose that PU.1 mutants Lys240Arg and Tyr244Ala have higher potential for myelopoiesis and this could be attributed to their reduced binding with the repressor protein GATA-1. Further studies such as colony formation assays will define more precisely the mechanism underlying PU.1 and GATA-1 antagonism as well as enhanced myeloid lineage formation ability of PU.1 mutants and ultimately may lead to the development of approaches for better differentiation therapy.

ACKNOWLEDGMENTS

We thank Col. Velu Nair, Army Hospital (R&R), Delhi for providing human bone marrow samples. We also thank Dr. B.S. Dwarkanath and Dr. J.S. Adhikari for flow cytometric analyses. Pallavi Gupta in particular thanks Indian Council of Medical Research (ICMR) for the award of Research Fellowship. This work was funded by DRDO, Govt. of India.

REFERENCES

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; **215**: 403-410.
2. Arnold K, Bordoli L, Kopp J, Schwede, T. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 2006; **22**: 195-201.
3. Behre G, Whitmarsh AJ, Coghlan MP, Hoang T, Carpenter CL, Zhang DE, *et al.* c-Jun is a JNK-independent coactivator of the PU.1 transcription factor. J Biol Chem 1999; **274**: 4939-4946.
4. Cavallo L, Kleinjung J, Fraternali F. POPS: A fast algorithm for solvent accessible surface areas at atomic and residue level. *Nucleic Acids Res* 2003; **31**: 3364-3366.
5. Gossen M, Bonin AL, Freundlieb S, Bujard H. Inducible gene expression systems for higher eukaryotic cells. *Curr Opin Biotechnol* 1994; **5**: 516-520.

6. Gossen M, Freundlieb S, Bender G, Muller G, Hillen W, Bujard H. Transcriptional activation by tetracycline in mammalian cells. *Science* 1995; **268**: 1766-1769.
7. Higgins DG, Thompson JD, Gibson TJ. Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 1996; **266**: 382-402.
8. Hooft RWW, Vriend G, Sander C, Abola EE. Errors in protein structures. *Nature* 1996; **381**: 272-272.
9. Klemsz M, Maki R. Activation of transcription by PU.1 requires both acidic and glutamine domains. *Mol Cell Biol* 1996; **16**: 390-397.
10. Klemsz MJ, McKercher SR, Celada A, Beveren CV, Maki RA. The macrophage and B cell specific transcription factor PU.1 is related to the ets oncogene. *Cell* 1990; **61**: 113-124.
11. Kodandapani R, Pio F, Ni CZ, Piccialli G, Klemsz, M, McKercher S, *et al.* A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. *Nature* 1996; **380**: 456-460.
12. Nerlov C, Querfurth E, Kulessa H, Graf T. GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. *Blood* 2000; **95**: 2543-2551.
13. Neuvirth H, Heinemann U, Birnbaum D, Tishby N, Schreiber G. ProMateus-an open research approach to protein-binding sites analysis. *Nucleic Acids Res* 2007; **35**: W543-W548.
14. Ray D, Bosselut R, Ghysdael J, Mattei MG, Tavitian A, Moreau-Gachelin F. Characterization of Spi-B, a transcription factor related to the putative oncoprotein Spi-1/PU.1. *Mol Cell Biol* 1992; **12**: 4297-4304.
15. Rekhtman N, Radparvar F, Evans T, Skoultchi AI. Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. *Genes Dev* 1999; **13**: 1398-1411.
16. Ritchie DW. Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2. *Proteins Struct Func Genet* 2003; **52**: 98-106.
17. Ritchie DW, Kemp GJL. Protein docking using spherical polar Fourier correlations. *Proteins Struct Func Genet* 2000; **39**: 178-194.
18. Rogers S, Wells R, Rechsteiner M. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* 1986; **234**: 364-368.
19. Rosenbauer F, Wagner K, Kutok JL, Iwasaki H, Le Beau MM, Okuno Y, *et al.* Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat Genet* 2004; **36**: 624-630.
20. Schneider U, Schwenk HU, Bornkamm G. Characterization of EBV-genome negative "null" and "T" cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin lymphoma. *Int J Cancer* 1977; **19**: 621-626.

21. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* 2003; **31**: 3381-3385.

22. Spain LM, Guerriero A, Kunjibettu S, Scott EW. T-cell development in PU.1-deficient mice. *J Immunol* 1999; **163**: 2681-2687.

23. Tjandra N, Omichinski JG, Gronenborn AM, Clore GM, Bax A. Use of dipolar 1H-15N and 1H-13C couplings in the structure determination of magnetically oriented macromolecules in solution. Nat Struc Biol 1997 **4**; 732-738.

24. Visvader JE, Elefanty AG, Strasser A, Adam JM. GATA-1 but not SCL induces megakaryocytic differentiation in an early myeloid line. *EMBO J* 1992; **11**: 4557-4564.

25. Vriend G. WHAT IF: A molecular modeling and drug design program. *J Mol Graph* 1990; **8**, 52-56.

26. Zhang P, Behre G, Pan J, Iwama A, Auron PE, Tenen DG, and Sun Z. Negative cross-talk between hematopoietic regulators: GATA-1 proteins repress PU.1. *Proc Natl Acad Sci USA* 1999; **96**: 8705-8710.

FIGURES

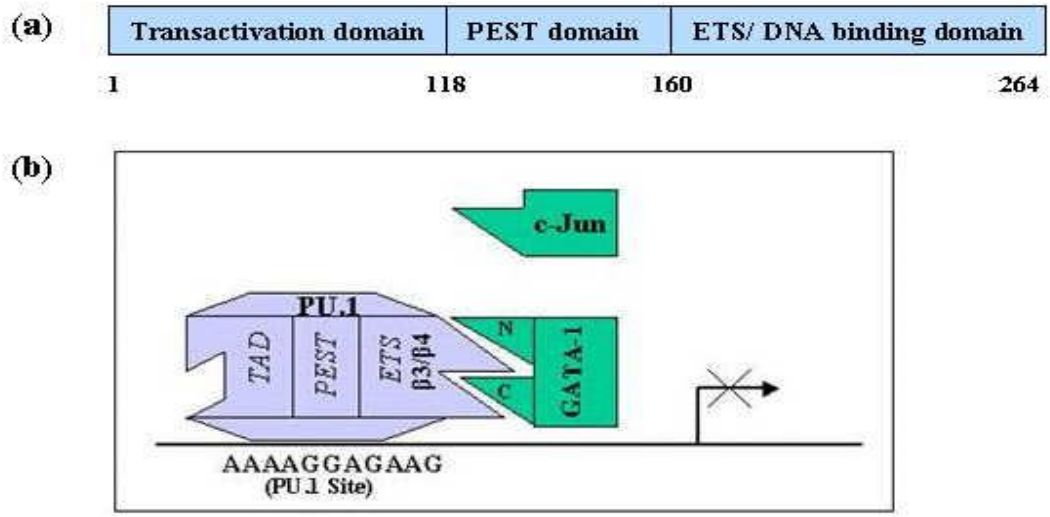

Figure 1

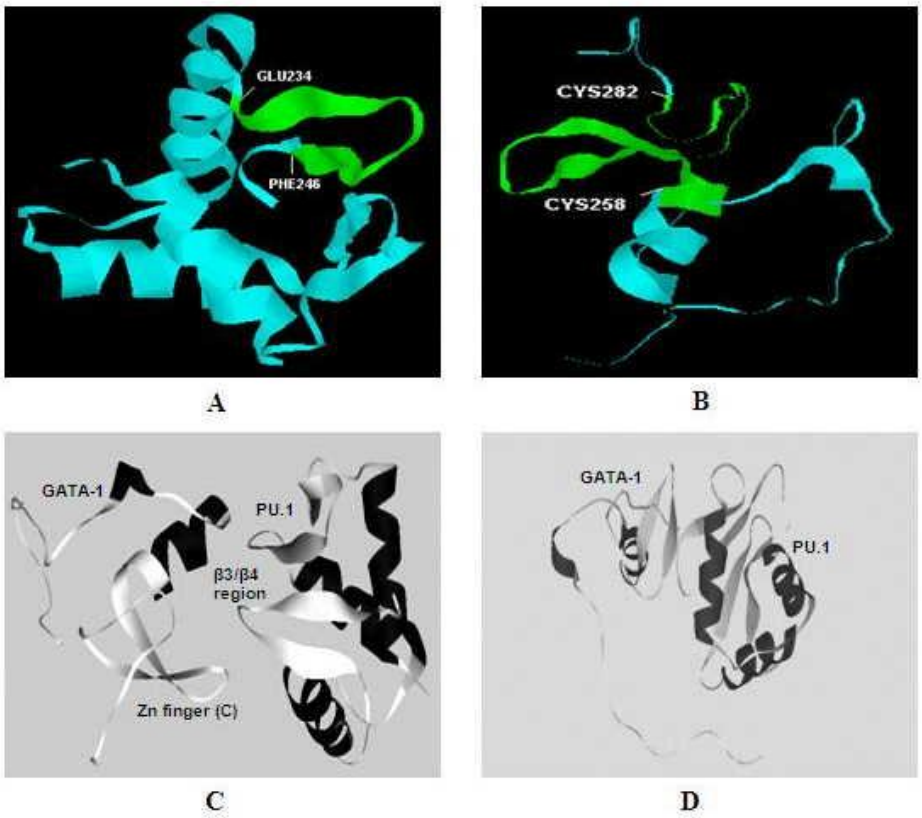

Figure 2

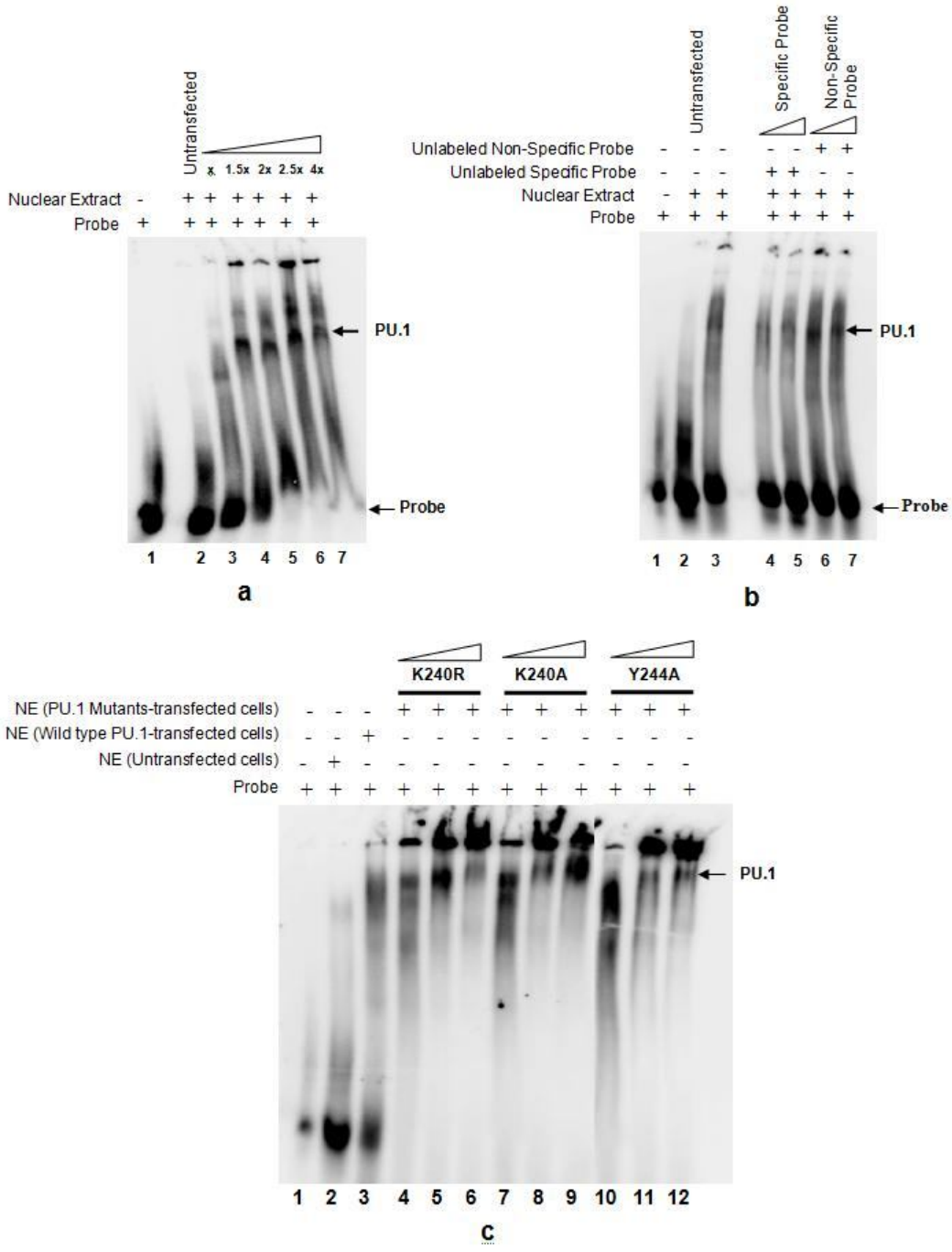

Figure 3

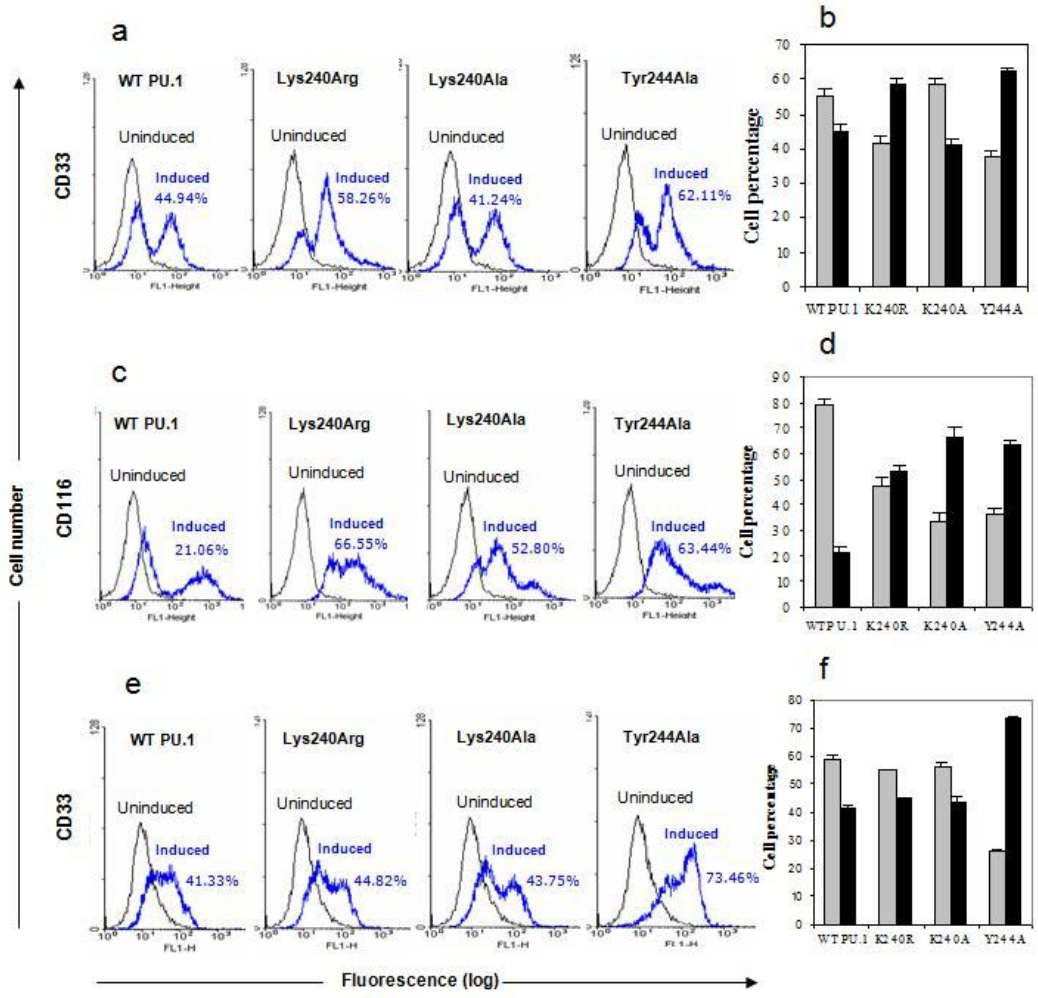

Figure 4

FIGURE LEGENDS

Figure 1. (a) Schematic diagram of PU.1. PU.1 has three functional domains, the amino terminal transactivation domain (TAD) (amino acids 1-86), the PEST domain (118-160) and the carboxy terminal ETS DNA binding domain (160-264). **(b) GATA-1 inhibits PU.1 function.** The C-terminal Zn finger of GATA-1 interacts with the β3/β4 region of PU.1, and displaces its coactivator c-Jun, thus repressing PU.1's function in a dose-dependent manner.

Figure. 2. Generation of 3D models of PU.1 and GATA-1 and their protein-protein docking. (a) 3D model of PU.1 DNA binding domain (residues 163-252) (based on the X-ray diffraction coordinates of PU.1-DNA complex; protein data bank code,

1pueF) **(Kodandpani et al., 1996)** was generated through SWISS-MODEL. The β3/β4 region (residues 234-246) has been highlighted in green color. **(b)** 3D model of GATA-1 (residues 252-316) (based on the NMR spectroscopic data of GATA-1-DNA complex; protein data bank code, gat1) **(Tjandra et al., 1997)** was generated by SWISS-MODEL. The C-terminal zinc finger of GATA-1 (residues 252-282) has been colored green. **(c)** Protein-protein complex between wild type PU.1 and GATA-1 as obtained by protein-protein docking between the two proteins using software Hex 4.5 **(Ritchie, 2000, 2003)**. The β3/β4 region of PU.1 interacts with the C-terminal zinc finger of GATA-1. The expected ligand-binding site PU.1 was positioned close to the c-terminal zinc finger residues of GATA-1 and docking was performed with receptor cutoff angle of 180°. **(d)** Protein-protein docking between Tyr244Ala mutant and GATA-1 performed by using Hex 4.5 software. The β3/β4 region of PU.1 does not interact with the C-terminal zinc finger of GATA-1.

Figure 3. **DNA binding ability of PU.1 mutants**. EMSA performed with nuclear extracts (NE) of Jurkat E6.1 cells transfected with wild type (WT) PU.1. Double stranded probes were end-labeled with [γ-^{32}P] dATP (BRIT, India) at a concentration of 10 mCi/reaction and purified using QIAGEN spin columns. The DNA binding reaction was performed at 4 °C for 1 hr in a total volume of 30 µl containing 5X binding buffer (1 M HEPES (pH 8.0), 250 mM KCl, 2.5 mM DTT, 0.25 mM EDTA, 5 mM $MgCl_2$, 25% Glycerol). The protein-DNA complexes were resolved in non-denaturing 4% polyacrylamide gel containing 2% glycerol in TBE buffer (1X: 89mM Tris-base, 89mM Boric acid, 2mM EDTA). Autoradiography was performed on dry gels using Phosphor Imager (Amersham Biosciences). **(a)** Increasing amounts of NE from cells expressing WT PU.1 were incubated with labeled probe (Lanes 4-8). PU.1-DNA binding increased with increasing amounts of NE. **(b)** Competition assay with unlabeled specific and non-specific probes (1:2 and 1:10 concentrations). Unlabeled specific probe competed with labeled probe; therefore no shift was observed (lanes 5 and 6). Unlabeled non-specific probe failed to compete with labeled probe; therefore a shift was observed (lanes 7 and 8). **(c)** EMSA performed with NE of transfectants overexpressing PU.1 mutants. Increasing concentrations of NE of mutants Lys240Arg (K240R), Lys240Ala (K240A) and Tyr244Ala (Y244A) were incubated with labeled probe. All the mutants were able to bind DNA, with increase in DNA binding ability upon increasing protein concentration. (Arrows on the right side of both the panels indicate the relative positions of PU.1-DNA complexes; x=15µg).

Oligonucleotide having a nucleotide exchange within the recognition region (PU.1 binding site underlined with mutated site bold) was used as a non-specific probe for competition assay. Primers: PU.1 Wt (sense): 5′-GCC TAG CTA AAA GGG GAA GAA GAG GAT CAG A-3′ ; PU.1 Wt (antisense): 5′-TCT GAT CCT CTT CTT CCC CTT TTA GCT AGG C-3′ ; Non-specific probe (sense): 5′-GCC TAG CTA AAA GG**TT**AA GAA GAG GAT CAG A-3′ ; Non-specific probe (antisense): 5′-TCT GAT CCT CTT AAC CTT TTA GCT AGG C-3′

Figure 4. Effects of overexpression of PU.1 and mutants in Jurkat E6.1 cells and HSCs. 1 x 10^6 transfected cells were washed twice with PBS and permeabilized using 1 ml of permeabilization buffer (0.5% (v/v) Tween in PBS) for 15 min and subsequently washed with staining buffer (2% FCS, 0.1% BSA and 0.1% sodium azide in PBS). The cells were then stained with the primary antibody for 45 min at 4 °C. The cells were washed 2-3 times with staining buffer and were further incubated with the fluorochrome-conjugated secondary antibody for 45 min at 4 °C. The cells were again washed 2-3 times with the staining buffer and were fixed with the 0.2% paraformaldehyde at 4 °C for 20 min and finally analysed under the flow cytometer FACS Caliber (BD Biosciences). FITC-conjugated mouse mono anti-human CD33 was purchased from Koma Biotech Inc. Mouse anti-human CD3 FITC and mouse anti-human CD116 (GM-CSFRα) FITC were procured from BD Pharmingen. The flow cytometric data was analysed using WinMDI 2.8 (Windows Multiple Document Interface). **(a,b,c,d)** Levels of CD33 and CD116 expression as measured by flow cytometry after 72 h of doxycycline induction. The levels of myeloid differentiation are shown in (a and c) as percentage of cells expressing CD33 and CD116 respectively alongside histograms and in (b and d) as mean fluorescence of cells expressing CD33 and CD116 respectively in each cell population. An increase in expression of CD33 and CD116 was found in Lys240Arg and Tyr244Ala transfectants. The maximum increase in CD33 expression was found in case of mutants Tyr244Ala (17% more than wild type PU.1) (P=0.05) and Lys240Arg (13%) (P=0.02) as compared to the wild type PU.1. An increase was found in CD116 expression in cells overexpressing mutants Lys240Arg and Tyr244Ala by 45% (P<0.01) and 42% (P<0.01) respectively as compared to wild type PU.1. (P<0.5 is taken as significant) **(e,f)** Level of CD33 expression in bone marrow-derived $CD34^+$ HSCs transfected with WT PU.1 and mutants after 96 h of doxycycline induction. An increase in CD33 expression was found in HSCs transfected with PU.1 mutant Tyr244Ala as shown in (e) percentage of cells expressing CD33 alongside histograms and in (f) as bar graph showing mean fluorescence of cells expressing CD33 in each cell population. A significant (32%) (P<0.01) increase was observed in CD33 expression in case of HSCs overexpressing mutant Tyr244Ala. The above result represents the data obtained from two individual experiments. (Grey bars represent cells expressing CD33/CD116; Black bars represent cells not expressing CD33/CD116).

Table 1: Validation of 3D models of PU.1 and GATA-1 through WHAT IF

Validity Checks Executed	Expected value	Observed value for PU.1	Observed value for GATA-1
Missing atoms, PU.1	0	Not detected	Not detected
Ramachandran Z-score	-4 to +4	-2.563 (OK)	-4.174 (very low)
All Bond Lengths	1.0	0.670 (OK)	0.626 (tight)
Bond length variability	--	0.013 (Normal)	0.013 (Low)
Bond Angles	1.0	0.907 (OK)	1.018
Bond angle variability	--	1.726 (Normal)	1.887 (Normal)
Chirality	--	OK	Deviations
chi-1/chi-2 angle correlation Z-score	-4 to +4	-1.915 (OK)	-1.400 (OK)
Backbone conformation Z-score	-4 to +4	-0.801 (OK)	-3.686 (poor)
Ist Generation Packing quality	-4 to +4	-0.656 (OK)	-6.663 (bad)
IInd generation packing quality	-4 to +4	-1.065 (OK)	-3.575 (poor)
Inside/Outside distribution	<1.5	0.898	1.024 (Normal)

Table 2: Solvent accessible surface area (SASA) for PU.1 ETS domain residues

Residue	PU.1 alone	Complex I	Complex II	Complex III	Complex IV	Complex V
Lys 237	86.8	86.8	44.4 (48.84%)	86.8	68.4	63.1
Val 238	65.5	65.5	23.4 (64.27%)	65.5	38.2	50.1
Lys 239	173.5	173.5	56.0 (67.72%)	173.5	94.6	144.8
Lys 240	155.6	155.6	86.0 (44.73%)	155.6	95.0	149.1

Table 3: Solvent accessibility surface area for GATA-1 C-terminal zinc finger residues

Residue	GATA-1 alone	Complex I	Complex II	Complex III	Complex IV	Complex V
Lys 252	220.9	144.8	121.9 (44.81%)	163.9	172.1	220.9
Leu 268	71.6	24.0	37.5 (47.62%)	29.9	22.0	61.6
Trp 269	51.5	21.9	22.1 (57.08%)	26.5	24.1	41.2
Arg 270	81.7	38.1	38.1 (53.36%)	43.1	55.5	55.0
Arg 271	108.3	49.0	46.5 (57.06%)	54.3	82.8	71.3
Asn 272	59.0	33.7	34.1 (42.2%)	36.4	52.9	39.5

(Among five minimum energy complexes of PU.1 and GATA-1, the complex II showed maximum decrease in SASA value for PU.1 as well as GATA-1 upon complex formation. Percentage decrease in SASA values has been indicated in parentheses)

Table 4:Changes in intramolecular H-bonding upon mutating Tyr244 to Ala.

S. No.	Intramolecular H-bonds	Wild type Complex 2 Bond length (A^o)	Mutant Tyr244Ala *Bond length* (A^o)
1	TYR219:HH-TYR244:OH	2.03096	----
2	TYR244:H-PHE193:O	1.97839	2.91411
3	LYS240:HN-VAL23 8:O	1.60888	2.46963
4	ARG175:HH22-LEU251:OXT	---	2.10312