www.impactjournals.com/oncotarget/

Research Paper

Expression and functional characterization of CD33 transcript variants in human acute myeloid leukemia

George S. Laszlo¹, Kimberly H. Harrington¹, Chelsea J. Gudgeon¹, Mary E. Beddoe¹, Matthew P. Fitzgibbon², Rhonda E. Ries¹, Jatinder K. Lamba³, Martin W. McIntosh², Soheil Meshinchi^{1,4,5}, Roland B. Walter^{1,6,7}

¹Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

²Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

³Department of Pharmacotherapy and Translational Research College of Pharmacy, University of Florida, Gainesville, FL, USA

⁴Children's Oncology Group, Arcadia, CA, USA

⁵Department of Pediatrics, University of Washington, Seattle, WA, USA

⁶Department of Medicine, Division of Hematology, University of Washington, Seattle, WA, USA

⁷Department of Epidemiology, University of Washington, Seattle, WA, USA

Correspondence to: Roland B. Walter, email: rwalter@fredhutch.org

Keywords: acute myeloid leukemia, antigen, CD33, immunotherapy, splice variants

Received: April 28, 2016 **Accepted:** May 17, 2016 **Published:** May 27, 2016

ABSTRACT

With the demonstration of improved survival of some acute myeloid leukemia (AML) patients with the CD33 antibody-drug conjugate, gemtuzumab ozogamicin (GO), CD33 has been validated as a target for antigen-specific immunotherapy. Since previous studies identified a CD33 splice variant missing exon 2 (CD33^{ΔE2}) and, consequently, the immune-dominant membrane-distal V-set domain, we investigated the expression and functional characteristics of CD33 transcript variants in AML. In primary AML specimens, we not only found full-length CD33 (CD33^{FL}) and CD33^{ΔE2} but also corresponding variants containing an alternate exon 7 predicted to encode a CD33 protein lacking most of the intracellular domain (CD33^{E7a} and, not previously described, CD33^{ΔE2,E7a}) in almost all cases. In acute leukemia cell sublines engineered to express individual CD33 splice variants, all splice variants had endocytic properties. CD33^{FL} and CD33^{E7a} mediated similar degrees of GO cytotoxicity, whereas CD33^{ΔE2} and CD33^{AE2,E7a} could not serve as target for GO. Co-expression of CD33^{AE2} did not interfere with CD33^{FL} endocytosis and did not impact CD33^{FL}-mediated GO cytotoxicity. Together, our findings document a greater-than-previously thought complexity of CD33 expression in human AML. They identify CD33 variants that lack exon 2 and are not recognized by current CD33-directed therapeutics as potential target for future unconjugated or conjugated antibodies.

INTRODUCTION

CD33, a member of the sialic acid-binding immunoglobulin (Ig)-like lectin (Siglec) family, is a 67 kD single pass transmembrane glycoprotein with endocytic properties that is primarily expressed on normal multipotent myeloid precursors, unipotent colony-forming cells, and maturing granulocytes and monocytes [1]. CD33 is comprised of an amino-terminal variable (V)-set Ig-like domain mediating sialic acid binding followed by a C2-set Ig-like domain in its extracellular region, a transmembrane domain, and a cytoplasmic tail that contains 2 conserved tyrosine-based inhibitory signaling motifs. Upon phosphorylation, the latter provide docking sites for the recruitment and activation of the Src homology-2 (SH2) domain-containing tyrosine phosphatases SHP-1 and SHP-2 [1]. Increasing evidence suggests that CD33 and related Siglecs with inhibitory signaling motifs modulate inflammatory and immune responses through dampening of tyrosine kinase-driven signaling pathways [2]. For CD33, the precise physiological functions have remained unclear, but recently, genome-wide association studies have identified variants of CD33 as a major risk factor for Alzheimer's disease, suggesting an important role in neurodegeneration [2].

Restricted expression patterns render Siglecs appealing targets for antibody- and glycan-based therapeutics [3]. Consistent with its characteristic as a myeloid differentiation antigen, CD33 is expressed on at least a subset of malignant blasts in nearly all patients with acute myeloid leukemia (AML) and also, perhaps, in underlying leukemia stem cells in some cases [1, 4]. There has therefore been a long-standing interest in the development of CD33-directed unconjugated and conjugated monoclonal antibodies for the treatment of AML. To date, gemtuzumab ozogamicin (GO), a humanized CD33 antibody conjugated to a calicheamicin- γ , derivative, has had the most success in the clinic. Several randomized trials have demonstrated that the addition of GO to induction chemotherapy reduces the relapse risk and improves survival of adults with newly diagnosed AML and favorable- or intermediate-risk features [5]. These studies are complemented by a large Children's Oncology Group (COG) trial of > 1,000 pediatric patients in which GO reduced the relapse risk and improved event-free survival irrespective of the cytogenetic disease risk [6]. Although GO was withdrawn from the commercial market in most countries in 2010, these results have validated CD33 as therapeutic target in AML and have sparked renewed interest in CD33-directed immunotherapies [1].

UCSC genome browser and Ensembl databases contain 3 unique, validated CD33 transcript variants supported by at least one non-suspect mRNA sequence, NM 001772, NM 001082618 and NM 001177608. NM 001772 represents full-length CD33 (CD33^{FL}) consisting of 7 coding exons and a 364 amino acid protein [7]. NM 001082618 skips exon 2 resulting in CD33 with a deleted exon 2 (CD33 $^{\Delta E2}$) and a 237 amino acid protein [8-11]. NM 001177608 uses an alternate exon 7, exon 7a (CD33^{E7a}) [10], which is predicted to encode an early translational stop, resulting in a 310 amino acid protein with truncation of the 54 c-terminal amino acids of the intracellular domain relative to CD33^{FL}. The existence of CD33 variants lacking exon 2 is relevant for therapeutic targeting of CD33 since the V-set domain, which is encoded by exon 2, contains immune-dominant epitope(s) [8, 12] recognized by all currently clinically exploited CD33-directed therapeutics. Moreover, our previous studies in cell lines engineered to express mutant forms of CD33 demonstrated that CD33 endocytosis is largely limited and determined by the intracellular domain of CD33 and that changes in this domain significantly impact CD33 internalization [13-15]. It is therefore conceivable that transcript variants that lack most of the intracellular domain of CD33 differ in their endocytic properties from the wild-type protein. Here, we investigated to what degree CD33 transcript variants are found in AML, what the endocytic properties are of these variants, and whether their co-expression affects the ability of CD33^{FL} to serve as a target for therapeutic antibodies.

RESULTS

Expression of CD33 splice variants in primary AML specimens

To survey the pattern of CD33 splice variant expression in human AML, we took advantage of whole transcriptome RNA sequencing (RNAseq) data from 61 pre-treatment bone marrow and 7 peripheral blood specimens collected from pediatric patients with newly diagnosed AML. CD33 coding sequence splice junctions were filtered for junctions with 10% prevalence at a threshold of ≥ 10 reads per specimen, resulting in 8 observed splice junctions (E1/E2, E1/E3, E2/E3, E3/E4, E4/E5, E5/E6, E6/E7a, E6/E7b; Figure 1A). These 8 splice junctions were consistent with the 3 known and wellsupported protein coding CD33 transcripts (NM 001772, NM 001082618, and NM 001177608; Figure 1B). Of the 8 junctions, 6 were shared by 2 or more known CD33 variants, whereas 2 were unique to one of two variants, NM 001082618 (E1/E3; specific for CD33^{AE2}) and NM 001177608 (E6/E7a; specific for CD33^{E7a}). The latter 2 were found in 57/68 (85.3%) and 55/68 (80.9%) of the specimens (Figure 1C). The observed CD33 splice junctions in our RNAseq dataset also offered the possibility of a previously unknown fourth, distinct CD33 transcript that lacked exon 2 and used exon E7a instead of E7b (CD33^{AE2,E7a}).

To further study CD33 variants in human AML and validate the findings obtained with RNAseq, we designed transcript-specific primers and used RT-PCR to assess expression of the CD33^{FL}, CD33^{△E2}, CD33^{E7a}, and CD33^{AE2,E7a} transcripts in 29 pre-treatment specimens from adults with AML. Basic characteristics of these specimens are summarized in Table 1. In the majority of specimens (27/29 [93.1%]), amplicons corresponding to the expected sizes of all 4 CD33 transcripts were found, whereas in the remaining 2 specimens (6.9%; samples #14 and #21), only 3 transcripts (CD33^{FL}, CD33^{ΔE2}, and CD33^{E7a}) were detectable (Figure 2). Three unique amplicons per CD33 transcript were subjected to Sanger sequencing to confirm the lack of exon 2 in CD33^{Δ E2}, use of exon 7a in CD33^{E7a}, and lack of exon 2 combined with use of exon 7a in CD33 $^{\Delta E2,E7a}$; the latter sequence was subsequently deposited in GenBank (accession no. KX139463). In 7 cases, we had bone marrow specimens (containing AML) cells) and paired peripheral blood specimens (without morphologic evidence of AML cells) available that were obtained on the same day. In all cases, the pattern of CD33 isoform expression was very similar in the peripheral blood and the marrow specimen, i.e. all CD33 isoforms could be identified in non-blast cells.

-				
Α	Exon/exon junction	Transcripts represented by splice junction	AML specimens with splice junction (%)	AML total reads (reads/positive specimens)
	E1/E2	NM_001772, NM_001177608	57/68 (83.8)	1215 (21.3)
	E1/E3	NM 001082618, CD33 ^{ΔE2,E7a}	57/68 (83.8)	1064 (18.7)
	E2/E3	NM_001772, NM_001177608	65/68 (95.6)	2661 (40.9)
	E3/E4	all	65/68 (95.6)	2836 (43.6)
	E4/E5	all	68/68 (100)	3705 (54.5)
	E5/E6	all	66/68 (97.1)	3852 (58.4)
	E6/E7a	NM 001177608, CD33 ^{ΔE2,E7a}	55/68 (80.9)	439 (8.0)
	E6/E7b	NM_001772, NM_001177608	65/68 (95.6)	3468 (53.4)





Sample #	Age	Disease stage	Karyotype	Cytogenetic risk	FLT3/ ITD	NPM1	Source	Blasts (%)
1	68.90	Refractory	Hypodiploid with monosomy of chromosome 5, 11, 13, and 17	Adverse	Neg	Neg	PB	86.9
2	67.35	Relapse	Normal	Intermediate	Pos	Neg	BM	94.7
3	65.30	Relapse	Normal	Intermediate	Pos	Pos	BM	89.2
4	53.01	Diagnosis	Normal	Intermediate	Pos	Neg	PB	87.4
5	71.52	Relapse	Normal	Intermediate	Pos	Neg	BM	82.0
6	78.79	Diagnosis	Normal	Intermediate	ND	ND	PB	95.1
7	63.00	Diagnosis	46,XX,t(3;5)(q21;q22) [9]/46,XX[11]	Intermediate	Neg	Neg	PB	92.3
8	67.14	Relapse	Normal	Intermediate	Neg	Pos	BM	79.4
9	61.07	Diagnosis	Normal	Intermediate	Neg	Pos	BM	94.1
10	78.03	Diagnosis	46,XX,del(7)(q32) [2]/46,XX[19]	Adverse	Neg	Neg	BM	90.9
11	72.63	Diagnosis	47,XY,add(1),del(1) (p22),add(8),del(16)	Adverse	Neg	Neg	BM	83.4
12	79.30	Diagnosis	Normal	Intermediate	ND	ND	BM	65.1
13	79.98	Diagnosis	44,XX,-3,-5,add(5)(p15) or del(5)(p13),inv(7) (p22q11.2),+i(11)(q10),add(12) (p13),-17[cp20]	Adverse	ND	ND	BM	58.7
14	48.48	Refractory	Normal	Intermediate	Neg	ND	PB	89.4
15	37.99	Refractory	Normal	Intermediate	Neg	Pos	PB	92.5
16	23.92	Diagnosis	46,XX,inv(16)(p13q22) (17]/47,sdl,+22[3]	Favorable	Neg	Neg	PB	92.2
17	68.71	Diagnosis	Normal	Intermediate	Pos	Pos	PB	93.5
18	73.77	Diagnosis	91,XXXX,-5,del(6)(q21q25)[20]	Adverse	Neg	Neg	BM	86.5
19	36.16	Relapse	47,XY,+8[20]	Intermediate	Pos	Neg	BM	85.8
20	54.43	Diagnosis	46,XY,inv(16)(p13.1q22)[20]	Favorable	Neg	Neg	PB	92.5
21	64.67	Diagnosis	91,XXYY,del(5)(q23q31),del(6) (q15q23),-7,del(7)(q31) [10]/94~99,s1,+6,+8,+9,+10,+13, +14,+19,+19,+20,+22,+0~1mar[cp6]/46,XY[4]	Adverse	Pos	Neg	BM	38.6
22	66.51	Diagnosis	Normal	Intermediate	Pos	Pos	BM	93.2
23	51.43	Diagnosis	Normal	Intermediate	Pos	Pos	PB	86.1
24	69.10	Refractory	Normal	Intermediate	Pos	Pos	PB	87.1
25	52.13	Diagnosis	Normal	Intermediate	Neg	Neg	PB	70.2
26	40.16	Diagnosis	Normal	Intermediate	Pos	Pos	PB	75.2
27	40.56	Relapse	Normal	Intermediate	Pos	Pos	PB	90.3
28	55.69	Refractory	Normal	Intermediate	Pos	ND	PB	83.2
29	48.30	Refractory	46,XY,t(11;12)(q13;p13) [9]/46,XY[1]	Intermediate	Neg	Neg	PB	87.4

Table 1: Basic characteristics of primary AML specimens used to study expression of CD33 splice variants by RT-PCR

Abbreviations: BM, bone marrow; ND, not done; neg, negative; PB, peripheral blood; pos, positive.

Generation of cell line models to study CD33 splice variants

Having demonstrated the presence of $CD33^{AE2}$, CD33^{E7a}, and CD33^{AE2,E7a} mRNA in primary human AML cells, we then studied the functional characteristics of encoded proteins. Therefore, we generated cell line models of human acute leukemia in which we selectively expressed untagged or His-tagged CD33^{FL}, CD33^{AE2}, CD33^{E7a} or CD33^{AE2,E7a} via lentivirus-mediated gene transfer, using a series of human AML (KG-1a, ML-1, and OCI-AML3) and acute lymphoblastic leukemia (ALL; RCH-ACV, REH, and RS4;11) cell lines. We used tagged versions of the CD33 proteins to distinguish those from endogenously expressed CD33 in some of these cell lines and to be able to track the CD33^{\Delta E2} and CD33^{\Delta E2,E7a} variants. While previous studies suggested that one CD33 antibody (clone Him3-4) specifically recognizes the C₂-set Ig-like domain and could be used to detect CD33 variants that lack exon 2 [9], in our experiments Him3-4 only recognized CD33^{FL} and CD33^{E7a} but not CD33^{ΔE2} or $CD33^{\Delta E2,E7a}$ when expressed in virally-transduced ALL cell lines that are devoid of endogenous CD33 (Jurkat, RCH-ACV, REH, or RS4;11: Figure 3) and only recognized CD33^{FL} but not CD33^{ΔE2} when expressed in virally-transduced HEK293T cells (Figure 4). As shown in Figure 5, engineered sublines expressed high amounts of CD33^{FL} and CD33^{E7a} in all cell lines. Lentivirus-mediated gene transfer also resulted in readily detectable cell surface display of the CD33^{ΔE2} and CD33^{ΔE2,E7a} splice variants, although the level of expression varied considerably across cell line backgrounds.

Internalization of CD33 splice variants in human leukemia cell lines

As the first characteristic, we studied the endocytic properties of the CD33 splice variants when engaged by bivalent antibodies. To test whether the internalization of antibody-bound CD33 differed depending on the type of antibody used (i.e. His antibody vs. CD33 antibody), we compared the 120- and 240-minute internalization of CD33 in RCH-ACV and REH engineered to expressed untagged and His-tagged CD33^{FL}. The studies showed very comparable rates of antibody uptake when studied with CD33 antibody in cells expressing untagged and His-tagged CD33^{FL} or with His antibody in cells expressing His-tagged CD33^{FL}, indicating that the His-tag does not interfere with the natural internalization properties of CD33 (Figure 6). Overall, our experiments in the panel



Figure 2: Expression of CD33 splice variants in primary AML specimens. Total RNA was isolated from 29 primary AML specimens and assayed for expression of mRNA for CD33^{FL}, CD33^{AE2}, CD33^{E7a}, and CD33^{AE2,E7a}, using variant-specific primers. Amplicons were electrophoretically separated in 0.9% agarose gel and visualized with ethidium bromide. Shown are results from 6 representative samples; also shown are a positive control (OCI-AML3 cells), a negative control (-reverse transcriptase [-RT]), as well as a sizing ladder. Lane 1: CD33^{FL} (expected amplicon size: 980 bp); lane 2: CD33^{AE2} (expected amplicon size: 617 bp); lane 3: CD33^{E7a} (expected amplicon size: 965 bp); and lane 4: CD33^{AE2,E7a} (expected amplicon size: 604 bp).

of sublines of AML and ALL cells expressing Histagged CD33 spice variants demonstrated that each variant is internalized when bound by a bivalent His antibody (Figure 5). While, on average, the internalization properties of each variant appeared relatively similar, we noted some cell line-specific differences in the comparative uptake of individual CD33 splice variants. In particular, in ML-1 cells, internalization of the CD33^{E7a} variant proceeded more slowly than that of the other CD33 molecules. Since some previous data have suggested that CD33 might form dimers on the cell surface in its physiologic state [16], we considered the possibility that CD33^{AE2} could interfere with the functional properties of CD33^{FL}. To test this idea, we compared the internalization of endogenous CD33 when bound with CD33 antibody in ML-1, OCI-AML3, and TF-1 cells that were lentivirally-transduced to also express CD33^{AE2} with that of non-transduced parental cells. These studies showed that uptake of endogenous CD33 was unaffected by the forced expression of the CD33^{AE2} splice variant (Figure 7). Finally, we also assessed the degree to which cell surface CD33 expression levels decreased with continued exposure to antibody ("CD33 modulation"). In 24-hour assays, we found some differences between individual CD33 variants in several acute leukemia cell lines; however, while highly reproducible, these



Figure 3: Phenotyping of acute leukemia cell lines with CD33 antibodies. Parental RCH-ACV, REH, and RS4;11 cells and their sublines transduced with His-tagged CD33^{FL}, CD33^{AE2}, CD33^{E7a}, or CD33^{AE2,E7a} were incubated with HIM3-4, P67.6 or isotype control antibody followed by incubation with a biotin-anti-mouse secondary antibody and APC-streptavidin. Results are shown from one representative experiment.



Figure 4: Phenotyping of engineered HEK293T cells with CD33 antibodies. Parental HEK293T cells and sublines transduced with untagged or His-tagged CD33^{FL} or CD33^{AE2} were incubated with HIM3-4, P67.6, a His antibody, or isotype control antibody followed by incubation with a biotin-anti-mouse secondary antibody and APC-streptavidin. Results are shown from one representative experiment.



Figure 5: Internalization of CD33 splice variants in engineered acute leukemia cell lines. Parental (A) KG-1a, (B) ML-1, (C) OCI-AML3, (D) RCH-ACV, (E) REH, and (F) RS4;11 cells were transduced with lentiviral particles to express either His-tagged CD33^{FL}, CD33^{AE2}, CD33^{E7a}, or CD33^{AE2,E7a}. *LEFT PANEL*: Cell surface display of the tagged transgenes was then quantified flow cyometrically using His antibody, and results expressed as arbitrary median fluorescence units (MFI). *RIGHT PANEL*: Internalization of His antibody was quantified in aliquots of engineered leukemia cell sublines over 120 or 240 minutes, and shown as percentage of remaining cell surface expression of the CD33 transgene detected at time 0. Results are depicted as mean ± SEM from 3 independent experiments.

differences were not consistent across cell lines but, rather, appeared cell context specific. Specifically, in engineered AML cell lines, a reduced degree of antigen modulation relative to wild-type CD33 was noted in sublines expressing CD33 isoforms containing exon 7a; by comparison, in engineered ALL cell lines, a reduced degree of antigen modulation relative to wild-type CD33 was noted in sublines expressing CD33 isoforms that lack exon 2 (Figure 8).

GO-induced cytotoxicity in human leukemia cell lines expressing CD33 splice variants

As the second characteristic, we studied the cytotoxic properties of the CD33 antibody-drug conjugate GO in parental AML and ALL cell lines and corresponding sublines engineered to express His-tagged CD33 splice variants. As shown in Figure 9, expression of CD33 splice variants that contain the V-set domain recognized by GO



Figure 6: Internalization of antibody-bound CD33. Comparative analysis of the 120-minute and 240-minute internalization of untagged and His-tagged $CD33^{FL}$ when bound with CD33 antibody or His antibody in lentivirally-transduced (A) REC-ACV and (B) REH cells. Results are shown as mean \pm SEM from 3 independent experiments.



Figure 7: Effect of CD33^{Δ E2} **on internalization of CD33**^{FL}**.** Parental (A) ML-1, (B) OCI-AML3, and (C) TF-1 cells and their sublines transduced with His-tagged CD33^{Δ E2} were incubated with CD33 antibody, and internalization of antibody-bound CD33 was then quantified flow cyometrically over 240 minutes and shown as percentage of remaining cell surface expression of CD33 detected at time 0. Results are depicted as mean ± SEM from 3 independent experiments.

(CD33^{FL} and CD33^{E7a}) sensitized human AML and ALL cell lines to GO-induced cytotoxicity. Consistent with our previous data [13], overexpression of CD33^{FL} increased GO-induced cytotoxicity in AML cell lines (KG-1a, ML-1, and OCI-AML3) that endogenously express CD33. In contrast, expression of CD33 variants that do not contain the V-set domain (CD33^{AE2} and CD33^{AE2,E7a}) did not modulate the cytotoxic effects of GO. Moreover, co-expression of CD33^{AE2} did not impact GO-induced cytotoxicity in AML cells endogenously displaying CD33 on the cell surface, as evidenced by our findings in ML-1 and OCI-AML3 cells, indicating that CD33^{AE2} does not interfere with the ability of CD33^{FL} to serve as carrier for intracellular delivery of cytotoxic drugs to CD33expressing leukemia cells.

DISCUSSION

number of antigens exploited The for immunotherapy of AML has been increasing rapidly over the last several years [17]. Nonetheless, CD33 remains of great interest for the development of antibody-based therapies. Besides its broad expression in AML, this interest is at least partially due to two circumstances, namely the fact that it thus far represents the only validated target in AML and the fact that GO - the antibody-drug conjugate that provided target validation - is currently not commercially available in most countries [1]. Several new CD33-directed agents, including a second-generation antibody-drug conjugate (SGN-CD33A), a bispecific T-cell engaging (BiTE) antibody (AMG 330), and chimeric antigen receptormodified T-cells have recently entered the clinical phase of drug testing, and others are at various stages of preclinical development.

Because of this ongoing interest in CD33 as an AML target, we sought to better characterize the expression pattern of various forms of CD33 in human AML. Using a large set of diagnostic specimens from patients with AML, our studies demonstrate that in the vast majority of cases, 3 splice variants of CD33 are found in addition to CD33^{FL}. One of these variants (CD33^{ΔE2}) has previously been found in normal myeloid cells, microglial cells, T-lymphocytes, and natural killer (NK) cells [8-11, 18], and has been implicated in influencing the susceptibility to Alzheimer's disease [10, 11, 18]. CD33 $^{\Delta E2}$ has also been described in several human AML cell lines [9]. We extend these findings by demonstrating that CD33^{ΔE2} is present at the mRNA level in the vast majority of primary AML specimens (29 of 29 in our series). Our studies also show the almost universal presence of mRNA corresponding to a CD33^{E7a} splice variant (the sequence of which was previously deposited in the genome databases; found in 29 of 29 in our series) and a CD33^{AE2,E7a} splice variant (found in 27 of 29 in our series), which has so far not been described.

Exon 2 encodes the ligand-binding extracellular V-set domain of CD33, and CD33 proteins lacking this exon are expected to feature functional properties that differ from those of wild-type CD33. Indeed, previous studies have indicated that CD33^{FL} but not CD33^{ΔE2}



Figure 8: CD33 modulation in engineered acute leukemia cells. Comparative analysis of the modulation of His-tagged CD33^{FL}, CD33^{ΔE2}, CD33^{ΔE2}, CD33^{ΔE2,E7a} in lentivirally-transduced AML cells (ML-1, OCI-AML3) and ALL cells (RCH-ACV, REH, RS4;11). Cells were incubated for 24 hours with His antibody followed by incubation with a biotin-anti-mouse secondary antibody and APC-streptavidin. Results are shown as mean \pm SEM from 3 independent experiments.

inhibits uptake and clearance of insoluble, toxic amyloid β species by microglial cells *in vitro* – an activity that may be important for the pathogenesis of Alzheimer's disease [19]. These studies have identified CD33 as a potential target for the treatment and/or prevention of Alzheimer's disease [19]. In this situation, blocking CD33 isoforms that contain exon 2 and have functional activity may be sufficient. For the treatment of CD33+ malignancies, however, directing the therapeutic toward all CD33 isoforms displayed on the cell surface may be advantageous since it would provide the greatest possible abundance of antibody binding sites. Consistent with previous reports [9, 19], we found that CD33 variants lacking exon 2 can be expressed on the cell surface, although our studies indicate that the efficiency with which this happens varies between different cell line backgrounds. Similar to wild-type CD33, our experiments in engineered acute leukemia cell lines document that the CD33^{AE2} variant is internalized, and could thus also serve as target for CD33-directed therapeutics that depend on intracellular delivery of a toxic payload.

Our studies in lentivirally-transduced acute leukemia cell lines also show that the CD33 variants that contain exon 7a, and therefore lack almost the entire cytoplasmic tail of CD33, are internalized. At first glance, this is surprising since we previously found CD33 endocytosis to be controlled by the intracellular domain of CD33. Introduction of point mutations in this domain, for example in the immunoreceptor tyrosinebased inhibitory motifs or clusters of lysine residues, reduced internalization of antibody/CD33 complexes in our earlier studies [13, 15]. Whether the cell membrane localization is influenced by the presence of the cytoplasmic tail of CD33, and whether the mechanistic principles for the uptake process differ between wildtype CD33 and variants that contain exon 7a, is currently unknown and will be subject of future investigations. The cell context-specific differences we found with regard to modulation of individual CD33 variants would be consistent with such differences in membrane localization and/or internalization mechanisms. Planned studies will also aim to identify differences between CD33 variants containing exon 7a with wild-type protein with regard to suppression of myeloid cell function - we hypothesize that such differences exist given the lack of immunoreceptor tyrosine-based inhibitory motifs when exon 7a is utilized.

In summary, our studies demonstrate the presence of 3 splice variants of CD33 that have endocytic properties when bound by a bivalent antibody in almost all patients with AML. These findings document a greater-thanpreviously thought complexity of CD33 expression in human AML. Furthermore, they identify CD33 variants that lack exon 2 and are not recognized by currently explored CD33-directed therapeutics as potential, hitherto unexploited, targets for immunotherapy with unconjugated or conjugated CD33 antibodies.



Figure 9: Effect of CD33 splice variants on GO-induced cytotoxicity. Parental (A) KG-1a, (B) ML-1, (C) OCI-AML3, (D) RCH-ACV, (E) REH, and (F) RS4;11 cells and their sublines transduced with His-tagged CD33^{FL}, CD33^{AE2}, CD33^{E7a}, or CD33^{AE2,E7a} were then incubated with various concentrations of GO for 3 days before cytotoxicity was assessed with DAPI staining. Increases in the percentage of DAPI⁺ cells in GO-treated cells are compared to corresponding cells that were incubated without GO, and results are shown as mean \pm SEM from 3 independent experiments performed in duplicate wells.

MATERIALS AND METHODS

Transcriptome sequencing of primary AML specimens

Sixty-eight patients with newly diagnosed AML enrolled on 2 recent COG trials were selected for retrospective whole transcriptome RNA sequencing (RNAseq) because they lacked known high-risk cytogenetic features but eventually relapsed [20]. Total RNA from pre-treatment bone marrow or peripheral blood specimens was used to generate a cDNA library, which was purified and enriched by polymerase chain reaction (PCR) amplification and subjected to 50-cycle pairedend sequencing on the Illumina HiSeq as previously described [20]. RNA sequencing reads were then aligned to the human genome to identify splice junctions between exons and identify new gene variants using TopHat [21]. Informed consent was obtained in accordance with the Declaration of Helsinki. The institutional review boards (IRBs) of all participating institutions approved the clinical protocol, while the Fred Hutchinson Cancer Research Center (Fred Hutch) IRB and the COG Myeloid Disease Biology Committee approved this research study.

PCR-based detection of CD33 splice variant expression in primary AML specimens

Frozen aliquots of Ficoll-isolated mononuclear cells from pretreatment specimens (bone marrow [n = 13], peripheral blood [n = 14], or apheresis [n = 2]) from 29 adults with newly diagnosed (n = 17) or relapsed/ refractory (n = 12) AML were obtained from Fred Hutch AML cell repositories and selected for inclusion based on the presence of a large number of myeloblasts (median: 87.4%; range: 38.6-95.1%; Table 1). We used the refined United Kingdom Medical Research Council/National Cancer Research Institute (MRC/NCRI) criteria to assign cytogenetic risk [22]. Patients provided written informed consent for the collection and use of their specimens for research purposes under an IRB-approved protocol. Genetic material was extracted using AllPrep DNA/ RNA Mini Kits (Qiagen, Valencia, CA). Approximately 25 ng of RNA was then reverse transcribed for each PCR reaction with specific primers to amplify CD33^{FL}, CD33^{AE2}, CD33^{E7a}, and CD33^{AE2,E7a} (see Table 2 for primer sequences). Amplicons were electrophoretically separated in 0.9% agarose gel and visualized with ethidium bromide (AlphaImager HP; ProteinSimple, San Jose, CA, USA). Selected amplicons were subjected to Applied Biosystems (ABI) BigDyeTM Terminator sequencing.

Generation of lentiviral vectors expressing CD33 constructs

A pRRLsin.cPPT.MSCV lentivirus containing a human CD33 $^{\rm FL}$ in conjunction with an internal ribosomal

entry site (IRES)/Enhanced Green Fluorescent Protein (EGFP) cassette has previously been described [13, 23]. Human cDNAs corresponding to CD33^{FL} with an N-terminal hexahistidine (His), and the CD33^{ΔE2}, CD33^{E7a}, and CD33^{ΔE2,E7a} variants with or without an N-terminal His tag were generated in TOPO vectors via standard PCR cloning procedures (see Table 3 for primer sequences), verified by sequencing, and subsequently transferred into the pRRLsin.cPPT.MSCV vector. Lentiviral particles were prepared as described previously [13].

Parental and engineered human cell lines

Human myeloid KG-1a cells (kindly provided by Dr. Derek L. Stirewalt; Fred Hutch, Seattle, WA), OCI-AML3 cells (kindly provided by Dr. Soheil Meshinchi; Fred Hutch), and ML-1 as well as TF-1 cells (kindly provided by Dr. Irwin D. Bernstein; Fred Hutch) were maintained as previously described [13, 24, 25]. Human lymphoid RCH-ACV and REH cells (kindly provided by Dr. Jerald P. Radich; Fred Hutch) were maintained in RPMI-1640 medium (Life Technologies, Grand Island, NY) with 10% fetal bovine serum (FBS; HyClone, Thermo Scientific, Logan, UT), and RS4;11 cells were maintained in MEM Alpha (Life Technologies) with 10% FBS. Human embryo kidney (HEK)293T cells (kindly provided by Dr. Jonathan A. Cooper; Fred Hutch) were maintained in DMEM medium (Life Technologies) with 10% FBS. All cell lines were confirmed to be mycoplasma-free but were not recently authenticated. Sublines of these cell lines overexpressing individual tagged or untagged CD33 constructs were generated through transduction with the appropriate pRRLsin.cPPT.MSCV lentiviral particles at a multiplicity of infection (MOI) of 25. EGFP-positive cells were isolated by flow cytometry and re-cultured for further analysis.

Quantification of CD33 expression

Expression of untagged CD33 variants on leukemia cell lines was quantified by flow cytometry using an unconjugated CD33 antibody (clone P67.6; BD Biosciences, San Jose, CA, USA) [13, 23] or unconjugated CD33 antibody followed by a biotin-anti-mouse secondary antibody and APC-streptavidin (both BD Biosciences). In some cases, a second CD33 antibody (clone HIM3-4) previously reported to recognize the C2-set Ig-like domain of CD33 [9] was used to detect cell surface display of CD33 variants. Expression of hexahistidine-tagged CD33 proteins was determined using an antibody against THETM His tag (clone 6G2A9; GenScript, Piscataway, NJ, USA) followed by a biotin-anti-mouse secondary antibody and APC-streptavidin. To identify nonviable cells, samples were stained with 4',6-diamidino-2-phenylindole (DAPI). 10,000 events were acquired on a BD FACSCanto II flow cytometer (BD Biosciences), and DAPI⁻ cells analyzed using FlowJo (Tree Star, Ashland, OR).

CD33 variant	Primer sequences
CD33 ^{FL}	For: 5'-TGGCTATGGATCCAAATTTCTGGCTG-3' (annealing to exon 2) Rev: 5'-AGCATAATGCAGCTCCTCATC-3' (annealing to exon 7b)
CD33 ^{AE2}	For: 5'-CCTGCTGTGGGCAGACTTGAC-3' (annealing to junction of exon 1/exon 3) Rev: 5'-AGCATAATGCAGCTCCTCATC-3' (annealing to exon 7b)
CD33 ^{E7a}	For: 5'-TGGCTATGGATCCAAATTTCTGGCTG-3' (annealing to exon 2) Rev: 5'-GATTGCAGGTGTGAAACACTG-3' (annealing to exon 7a)
CD33 ^{AE2,E7a}	For: 5'-CCTGCTGTGGGCAGACTTGAC-3' (annealing to junction of exon 1/exon 3) Rev: 5'-GATTGCAGGTGTGAAACACTG-3' (annealing to exon 7a)

Table 3: List of mutagenesis primers

Construct	Primer sequences
CD33 ^{FL/His}	For: 5'-CATCATCACCATCACCACGGAGGTGGAATGGATCCAAATTTCTGGCTGC-3' Rev: 5'-TCCACCTCCGTGGTGATGGTGATGATGAGCCAGGGCCCCTGC-3'
CD33 ^{AE2}	For: 5'-TGCCCCTGCTGTGGGCAGACTTGACCCACAGGCCC-3' Rev: 5'-GGGCCTGTGGGTCAAGTCTGCCCACAGCAGGGGCA-3'
CD33 ^{ΔE2/His}	For: 5'-CATCATCACCATCACCACGGAGGTGGACCCAAAATCCTCATCCCTG-3' Rev: 5'-TCCACCTCCGTGGTGATGGTGATGATGCCTGTGGGTCAAGTCTGC-3'
CD33 ^{E7a} , CD33 ^{E7a/His} , CD33 ^{ΔE2,E7a} , CD33 ^{ΔE2,E7a/His}	For: 5'-ATGCCGCTGCTGCTACTGCTGC-3' Rev: 5'-TCAACGTACCGGGGAGGCTGACCCTGTGGTAG-3'

Quantification of CD33 internalization and modulation

To measure internalization of antibody-bound CD33, cells (typically $1-1.5 \times 10^6$) were transferred into 15 mL polypropylene conical bottom tubes (FalconTM, Corning Life Sciences, Tewksbury, MA, USA) and incubated for at least 20 minutes with phosphate buffered saline (PBS, GIBCO by Life Technologies)/2% FBS containing 2.5 µg/mL unconjugated P67.6 (Santa Cruz Biotechnology, Dallas, TX, USA) or His tag antibody in ice-water (to prevent internalization during the staining procedure), as appropriate. Cells were then washed in ice-cold PBS, resuspended in IMDM medium without antibody, split into several tubes, and incubated at 37°C (in 5% CO₂ and air) for various periods of time. Afterwards, cells were chilled and incubated with biotin-conjugated rat anti-mouse IgG, monoclonal antibody (used at 2.5 µg/mL in PBS/2% FBS), followed by incubation with streptavidin-APC conjugate (used at 2.5 µg/mL in PBS/2% FBS: both from BD Biosciences, San Jose, CA, USA) to detect remaining primary antibody on the cell surface. One sample that was kept in ice water was used to determine the starting level of antibody bound to the cell surface. To determine CD33 modulation, aliquots of engineered leukemia cells were left untreated or incubated with unlabeled His tag antibody (used at 2.5 µg/mL). After 24 hours, cells were washed in ice-cold PBS to remove unbound antibody and resuspended in PBS/2%FBS. Aliquots of untreated and antibody-treated cells were then incubated with His antibody (to saturate binding sites) or no antibody followed by biotin-conjugated rat anti-mouse IgG_1 (used at 2.5 µg/mL in PBS/2% FBS) and streptavidin-APC (used at 2.5 µg/mL in PBS/2% FBS). For analysis, all samples were stained with DAPI to identify nonviable cells; at least 10,000 events were acquired, and DAPI⁻ cells were analyzed on a BD FACSCanto II flow cytometer using FlowJo. Linear fluorescence values were used to calculate the percentage of CD33 internalization or modulation.

Quantification of drug-induced cytotoxicity

AML cells were incubated at 37°C (in 5% CO₂ and air) in 96-well round bottom plates (FalconTM) at 8×10^3 cells/well in 200 µL culture medium containing various concentrations of gemtuzumab ozogamicin (commercially obtained from Pfizer, New York, NY). After 72 hours, cell numbers and drug-induced cytotoxicity, using DAPI to detect non-viable cells, were determined using a BDFACSCanto II flow cytometer and analyzed with FlowJo software.

ACKNOWLEDGMENTS

We thank the Fred Hutch Co-operative Center of Excellence in Hematology (CCEH) Vector Production core facility for preparing lentiviral particles and the COG AML Reference Laboratory for providing diagnostic AML specimens.

CONFLICTS OF INTEREST

R.B.W. has received research funding from Amgen, Inc., Amphivena Therapeutics, Inc., Covagen AG, Emergent Biosolutions, Inc., Pfizer, Inc., and Seattle Genetics, Inc and Seattle Genetics, Inc., and is a consultant for Amphivena Therapeutics, Inc, and Covagen AG. The other authors declare no competing financial interests.

GRANT SUPPORT

Supported by grants R21-CA1555524, R01-CA114563, U01-CA176270, and U10-CA180886 from the National Cancer Institute/National Institutes of Health (NCI/NIH), P30-DK056465 (CCEH) from the National Institute of Diabetes and Digestive and Kidney Diseases/NIH, an Innovation Grant from the Alex's Lemonade Stand Foundation, and a Translational Research Program grant from the Leukemia & Lymphoma Society (6489–16). R.B.W. is a Leukemia & Lymphoma Society Scholar in Clinical Research (2311–15).

REFERENCES

- Laszlo GS, Estey EH, Walter RB. The past and future of CD33 as therapeutic target in acute myeloid leukemia. Blood Rev. 2014; 28:143–153.
- Macauley MS, Crocker PR, Paulson JC. Siglec-mediated regulation of immune cell function in disease. Nat Rev Immunol. 2014; 14:653–666.
- 3. Angata T, Nycholat CM, Macauley MS. Therapeutic Targeting of Siglecs using Antibody- and Glycan-Based Approaches. Trends Pharmacol Sci. 2015; 36:645–660.
- 4. Walter RB, Appelbaum FR, Estey EH, Bernstein ID. Acute myeloid leukemia stem cells and CD33-targeted immunotherapy. Blood. 2012; 119:6198–6208.
- Hills RK, Castaigne S, Appelbaum FR, Delaunay J, Petersdorf S, Othus M, Estey EH, Dombret H, Chevret S, Ifrah N, Cahn JY, Recher C, Chilton L, et al. Addition of gemtuzumab ozogamicin to induction chemotherapy in adult patients with acute myeloid leukaemia: a meta-analysis of individual patient data from randomised controlled trials. The Lancet Oncology. 2014; 15:986–996.
- 6. Gamis AS, Alonzo TA, Meshinchi S, Sung L, Gerbing RB, Raimondi SC, Hirsch BA, Kahwash SB, Heerema-McKenney A, Winter L, Glick K, Davies SM, Byron P, et al. Gemtuzumab ozogamicin in children and adolescents with de novo acute myeloid leukemia improves eventfree survival by reducing relapse risk: results from the randomized phase III Children's Oncology Group trial AAML0531. J Clin Oncol. 2014; 32:3021–3032.
- Simmons D, Seed B. Isolation of a cDNA encoding CD33, a differentiation antigen of myeloid progenitor cells. J Immunol. 1988; 141:2797–2800.
- 8. Hernández-Caselles T, Martínez-Esparza M, Pérez-Oliva AB, Quintanilla-Cecconi AM, García-Alonso A,

Alvarez-López DM, García-Peñarrubia P. A study of CD33 (SIGLEC-3) antigen expression and function on activated human T and NK cells: two isoforms of CD33 are generated by alternative splicing. J Leukoc Biol. 2006; 79:46–58.

- Pérez-Oliva AB, Martínez-Esparza M, Vicente-Fernández JJ, Corral-San Miguel R, García-Peñarrubia P, Hernández-Caselles T. Epitope mapping, expression and post-translational modifications of two isoforms of CD33 (CD33M and CD33m) on lymphoid and myeloid human cells. Glycobiology. 2011; 21:757–770.
- Malik M, Simpson JF, Parikh I, Wilfred BR, Fardo DW, Nelson PT, Estus S. CD33 Alzheimer's risk-altering polymorphism, CD33 expression, and exon 2 splicing. J Neurosci. 2013; 33:13320–13325.
- Raj T, Ryan KJ, Replogle JM, Chibnik LB, Rosenkrantz L, Tang A, Rothamel K, Stranger BE, Bennett DA, Evans DA, De Jager PL, Bradshaw EM. CD33: increased inclusion of exon 2 implicates the Ig V-set domain in Alzheimer's disease susceptibility. Hum Mol Genet. 2014; 23:2729–2736.
- 12. Peiper SC, Leboeuf RD, Hughes CB, Prasthofer EF, Borowitz MJ, Dewutter-Dambuyant C, Katz DR, Walker WS, Ashmun RA, Look AT. Report on the CD33 cluster workshop: biochemical and genetic characterization of gp67. In: Knapp W, Dörken B, Gilks WR, Rieber EP, Schmidt RE, Stein H and von dem Borne EAGK, eds. Leukocyte Typing IV: White Cell Differentiation Antigens. (Oxford, UK: Oxford University Press). 814–816.
- Walter RB, Raden BW, Kamikura DM, Cooper JA, Bernstein ID. Influence of CD33 expression levels and ITIMdependent internalization on gemtuzumab ozogamicininduced cytotoxicity. Blood. 2005; 105:1295–1302.
- Walter RB, Raden BW, Zeng R, Häusermann P, Bernstein ID, Cooper JA. ITIM-dependent endocytosis of CD33-related Siglecs: role of intracellular domain, tyrosine phosphorylation, and the tyrosine phosphatases, Shp1 and Shp2. J Leukoc Biol. 2008; 83:200–211.
- Walter RB, Hausermann P, Raden BW, Teckchandani AM, Kamikura DM, Bernstein ID, Cooper JA. Phosphorylated ITIMs enable ubiquitylation of an inhibitory cell surface receptor. Traffic. 2008; 9:267–279.
- Paul SP, Taylor LS, Stansbury EK, McVicar DW. Myeloid specific human CD33 is an inhibitory receptor with differential ITIM function in recruiting the phosphatases SHP-1 and SHP-2. Blood. 2000; 96:483–490.
- Buckley SA, Walter RB. Antigen-specific immunotherapies for acute myeloid leukemia. Hematology Am Soc Hematol Educ Program. 2015; 2015:584–595.
- Schwarz F, Springer SA, Altheide TK, Varki NM, Gagneux P, Varki A. Human-specific derived alleles of CD33 and other genes protect against postreproductive cognitive decline. Proc Natl Acad Sci U S A. 2016; 113:74–79.
- Griciuc A, Serrano-Pozo A, Parrado AR, Lesinski AN, Asselin CN, Mullin K, Hooli B, Choi SH, Hyman BT, Tanzi RE. Alzheimer's disease risk gene CD33 inhibits

microglial uptake of amyloid beta. Neuron. 2013; 78:631-643.

- 20. Walter RB, Laszlo GS, Alonzo TA, Gerbing RB, Levy S, Fitzgibbon MP, Gudgeon CJ, Ries RE, Harrington KH, Raimondi SC, Hirsch BA, Gamis AS, M WM, et al. Significance of expression of ITGA5 and its splice variants in acute myeloid leukemia: a report from the Children's Oncology Group. Am J Hematol. 2013; 88:694–702.
- Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009; 25:1105–1111.
- 22. Grimwade D, Hills RK, Moorman AV, Walker H, Chatters S, Goldstone AH, Wheatley K, Harrison CJ, Burnett AK. Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. Blood. 2010; 116:354–365.
- Laszlo GS, Gudgeon CJ, Harrington KH, Dell'Aringa J, Newhall KJ, Means GD, Sinclair AM, Kischel R, Frankel SR, Walter RB. Cellular determinants for preclinical activity of a novel CD33/CD3 bispecific T-cell engager (BiTE) antibody, AMG 330, against human AML. Blood. 2014; 123:554–561.
- Walter RB, Raden BW, Hong TC, Flowers DA, Bernstein ID, Linenberger ML. Multidrug resistance protein attenuates gemtuzumab ozogamicin-induced cytotoxicity in acute myeloid leukemia cells. Blood. 2003; 102:1466–1473.
- 25. Walter RB, Raden BW, Cronk MR, Bernstein ID, Appelbaum FR, Banker DE. The peripheral benzodiazepine receptor ligand PK11195 overcomes different resistance mechanisms to sensitize AML cells to gemtuzumab ozogamicin. Blood. 2004; 103:4276–4284.