Cell and Molecular Determinants of In Vivo Efficacy of the BH3 Mimetic ABT-263 against Pediatric Acute Lymphoblastic Leukemia Xenografts

Santi Suryani1, Hernan Carol1, Triona Ni Chonghaile2, Viktoras Frismantas3, Chintanu Sarmah1, Laura High1, Beat Bornhauser3, Mark J. Cowley4, Barbara Szymanska1, Kathryn Evans1, Ingrid Boehm1, Elise Tonå1, Luke Jones1, Donya Moradi Manesh1, Raushan T. Kurmasheva5, Catherine Billups6, Warren Kaplan4, Anthony Letai2, Jean-Pierre Bourquin3, Peter J. Houghton5, Malcolm A. Smith7, and Richard B. Lock1

Abstract

Purpose: Predictive biomarkers are required to identify patients who may benefit from the use of BH3 mimetics such as ABT-263. This study investigated the efficacy of ABT-263 against a panel of patient-derived pediatric acute lymphoblastic leukemia (ALL) xenografts and utilized cell and molecular approaches to identify biomarkers that predict in vivo ABT-263 sensitivity.

Experimental Design: The in vivo efficacy of ABT-263 was tested against a panel of 31 patient-derived ALL xenografts composed of MLL-, BCP-, and T-ALL subtypes. Basal gene expression profiles of ALL xenografts were analyzed and confirmed by quantitative RT-PCR, protein expression and BH3 profiling. An in vitro coculture assay with immortalized human mesenchymal cells was utilized to build a predictive model of in vivo ABT-263 sensitivity.

Results: ABT-263 demonstrated impressive activity against pediatric ALL xenografts, with 19 of 31 achieving objective responses. Among BCL2 family members, in vivo ABT-263 sensitivity correlated best with low MCL1 mRNA expression levels. BH3 profiling revealed that resistance to ABT-263 correlated with mitochondrial priming by NOXA peptide, suggesting a functional role for MCL1 protein. Using an in vitro coculture assay, a predictive model of in vivo ABT-263 sensitivity was built. Testing this model against 11 xenografts predicted in vivo ABT-263 responses with high sensitivity (50%) and specificity (100%).

Conclusion: These results highlight the in vivo efficacy of ABT-263 against a broad range of pediatric ALL subtypes and shows that a combination of in vitro functional assays can be used to predict its in vivo efficacy.


Introduction

Despite significant improvements in the treatment of childhood acute lymphoblastic leukemia (ALL) over the past 5 decades, curing those patients who relapse with this most common pediatric malignancy remains a significant challenge (1). These relapse cases are often associated with broad-range drug resistance (2), which remains a significant problem, thus highlighting the need to develop new therapies. Because evasion of apoptosis is recognized as one of the hallmarks of cancer (3), recent drug development has focused on targeting key components of the apoptosis signaling pathway (4). The BCL2 family of proteins includes key regulators of the intrinsic apoptosis pathway, with cell fate being determined by the balance of prosurvival (e.g., BCL2 and MCL1) and proapoptotic (e.g., PUMA, NOXA) members (5, 6).

BH3-mimetic drugs, such as ABT-737 and its orally available analog ABT-263, were specifically designed to inhibit prosurvival BCL2 family proteins (7). Although these drugs bind with high affinity to BCL2, BCLW, and BCLXL, they exhibit lower affinity for MCL1 and A1 (7, 8). ABT-737 and ABT-263 have shown significant in vivo efficacy in preclinical xenograft models of hematolymphoid and solid malignancies (9, 10). Although clinical trials of ABT-263 in adults...
Manipulation of the apoptosis pathway is an appealing strategy for cancer treatment using BH3 mimetics such as ABT-263, although predictive biomarkers are required to identify patients who may benefit from their use. This study showed that ABT-263 exhibited broad in vivo efficacy against preclinical xenograft models of pediatric acute lymphoblastic leukemia (ALL). High MCL1 expression, at the mRNA and protein level, correlated with in vivo ABT-263 resistance, which was confirmed functionally by BH3 profiling. In addition, in vitro coculture assays predicted in vivo ABT-263 responses with high sensitivity and specificity. Therefore, a combination of functional assays could be used to predict ABT-263 activity in vivo. Given the strong efficacy of ABT-263 against a significant proportion of xenografts tested, these in-principle approaches could be included in the design of prospective clinical trials to determine if they can identify patients who may respond to treatment with this class of therapeutic agents.

have shown promising results (11–13), the main dose-limiting toxicity of thrombocytopenia has hindered its progression into pediatric patients. Consistent with the low affinity of ABT-737 and ABT-263 for MCL1, several reports have shown an inverse correlation of MCL1 expression with sensitivity to these drugs (14–16). Other proteins in the BCL2 family have also been implicated in determining sensitivity or resistance. For example, high BCL2 expression was associated with increased ABT-737 sensitivity in Non-Hodgkin’s lymphoma (NHL) cell lines and in murine fetal liver cells (15). However, recent studies provided evidence that MCL1 or prosurvival protein expression levels contribute to, but are not sufficient determinants of, resistance (17–20). Disruption of the interaction between MCL1 and BAK increased drug sensitivity (17, 18), suggesting that protein–protein interactions, rather than absolute levels, play a critical role in determining the sensitivity to BH3 mimetics. This interpretation was reinforced by “mitochondrial BH3 profiling,” which utilizes a panel of peptides derived from BH3-domains and their binding to antiapoptotic proteins, to predict a cell’s susceptibility to apoptosis induction (19, 20). Mitochondrial sensitivity to the BAD BH3 peptide, which has a pattern of interaction with antiapoptotic proteins similar to ABT-737 and ABT-263, was shown to predict in vitro ABT-737 sensitivity in small cell lung cancer, lymphoma, ALL, and acute myelogenous leukemia cell lines (19). Clinical responses to conventional chemotherapy in acute leukemia, multiple myeloma, and ovarian cancer instead were found to correlate with mitochondrial sensitivity with promiscuous interacting BH3 peptides such as Puma BH3 (20).

The Pediatric Preclinical Testing Program previously reported that ABT-263 was effective as a single agent against in vivo models of childhood cancer, and in particular pediatric ALL xenografts (10). The results suggested preferential efficacy against 2 T-cell ALL (T-ALL) in comparison to B-cell precursor (BCP)-ALL xenografts, albeit testing against a small panel of xenografts. In this study, we tested the in vivo efficacy of ABT-263 against a diverse panel of 31 molecularly characterized xenografts derived from T-ALL, BCP-ALL, and infant ALL with translocations of the mixed lineage leukemia (MLL) gene (infant MLL-ALL), as well as the efficacy of ABT-263 in combination with established drugs. To identify determinants of in vivo ABT-263 response, we then correlated gene expression profiles, mitochondrial BH3 profiling and in vitro ABT-263 sensitivity with single-agent ABT-263 efficacy. This powerful approach can be used as proof-of-principle to identify determinants of in vivo responses to other novel antileukemic drugs.

Materials and Methods

Xenografts and in vivo drug treatments

All experimental studies were conducted with approval from the Animal Care and Ethics Committee of the University of New South Wales (Sydney, Australia). Procedures by which we established continuous xenografts from childhood ALL biopsies in immune-deficient NOD/SCID (NOD.Cb17-Prkdcscid/SzJ) or NOD/SCID, IL2 receptor γ negative (NOD.Cg-Prkdcscid L2rgtm1Wjl/SzJ, NSG) mice, and tested their in vivo ABT-263 responses, have been described in detail previously (10, 21, 22). ALL subtypes were categorized at biopsy by their immunophenotype. Xenografts are available from the corresponding author upon request. ABT-263 (obtained from AbbVie under a standard Material Transfer Agreement) was administered orally at a dose of 100 mg/kg, daily for 21 days, as previously described (10). ABT-263 was also administered in combination with the conventional chemotherapeutic drugs vincristine (Baxter Healthcare; 1 mg/kg, days 0 and 7), dexamethasone (Sigma-Aldrich; 15 mg/kg, Mon–Fri 2 weeks) or L-asparaginase (Aventis; 1,500 U/kg, Mon–Fri 2 weeks) on a Mon to Fri × 2-week schedule at least 1 hour after administration of the established drug. It was necessary to attenuate the dose of ABT-263 to 25 mg/kg when combined with vincristine, and to 50 mg/kg when combined with dexamethasone and L-asparaginase in order to achieve a tolerable dose.

Leukemia engraftment and progression were assessed in groups of 6 to 10 female mice each of 20 to 25 g by weekly enumeration of the proportion of human CD45+ cells in the peripheral blood (%huCD45+; ref. 22). Individual mouse event-free survival (EFS) was calculated as the days from treatment initiation until the %huCD45+ reached 25%. EFS was represented graphically by Kaplan–Meier analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis. The efficacy of drug treatment was evaluated by leukemia growth delay (LGD), calculated as the difference between the median EFS of vehicle control and drug-treated groups of 6 to 10 female mice each of 20 to 25 g by weekly analysis.
expressed in a "COMPARE-like" format, which combines EFS and ORMs around the midpoint (0) representative of SD. Bars to the right or left of the midpoint represent objective responses or nonobjective responses, respectively. Significant and nonsignificant differences in EFS distribution between control and treated cohorts are represented by solid or dotted bars, respectively. Xenografts were excluded from analysis if >25% of mice within a cohort experienced nonleukemia-related toxicity or morbidity. Mice were excluded from the study if they developed spontaneous murine lymphomas.

To evaluate interactions between drugs in vivo, therapeutic enhancement was considered if the EFS of mice treated with the drug combination was significantly greater than those induced by both single agents used at their maximum tolerated doses (23, 24).

**Protein expression**

Preparation of extracts from xenograft cells previously harvested from the spleens of engrafted mice, determination of protein concentrations, and analysis of cellular proteins by immunoblotting have been described in detail elsewhere (25). Membranes were probed with anti-MCL1 (Genesearch) and anti-actin antibodies (Sigma-Aldrich) followed by horseradish peroxidase (HRP)—conjugated secondary antibody (GE Healthcare). Signal was detected by Immobilon Western Chemiluminescent HRP Substrate (Merck Millipore) and visualized using a VersaDoc 5000 Imaging System (Bio-Rad). Data were analyzed with QuantityOne software (Version 4.0; Bio-Rad).

**RNA extraction, real-time quantitative reverse transcription PCR and gene expression analysis**

Total RNA was extracted from xenograft cells, previously harvested from the spleens of engrafted mice and cryopreserved, using a combination of Trizol (Invitrogen) and Qiagen RNAeasy Kit. RNA was purified with QIAGEN RNAeasy spin columns, according to the manufacturer's protocol. RNA purity was considered acceptable if the ratio of OD260/280 was between 1.8 and 2.0. For use in microarrays, the RNA integrity number was determined using an Agilent Bioanalyzer and considered acceptable if >7.

Real-time quantitative reverse transcription PCR (RT-qPCR) was carried out using standard techniques. First-strand cDNA was synthesized using 2 µg of total RNA, random primers (Roche), and M-MLV Reverse Transcriptase (Invitrogen). Primers and probes for MCL1 were purchased from Life Technologies (Hs03043899_m1). Quantitative real-time PCR analysis was carried out in triplicate under the following cycling conditions: 50°C for 2 minutes and 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Elongation factor-1α (EF1α) was used as an internal normalization standard in each reaction (primers EF1aF, 5’-CTGGAACATCCAGGCCAAAT–3’; EF1aR, 5’-GCCGTGTGGCAATCCAAT-3’; probe, 5’-VIC-A GCCCGGGCFCATGCCCTG-TAMRA-3’).

RNA samples were used to prepare cRNA with Illumina TotalPrep RNA Amplification Kit (Life Technologies). cRNA was then hybridized to Illumina Human Beadchip HT12 Arrays. Gene expression datasets were analyzed using GenePattern v3.2.3 as we have previously described (26). Gene expression datasets can be accessed at www.ncbi.nlm.nih.gov/geo (Accession No. GSE52991; reviewer’s access: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=glw1gwusvaladfsr&acc=GSE52991). Benjamini and Hochberg false discovery rate (FDR; ref. 27) measurement and Smyth unadjusted P value (28) were used for evaluation of differential gene expression. Gene expression heatmaps were generated using GenePattern, whereby the range of color coding extends from minimum to maximum values per gene (per row). In each case, red indicates high, and blue low, level of expression. Unsupervised hierarchical clustering was performed using the Hierarchical Clustering module in GenePattern using the entire 47,323 probes representative of 34,694 genes present in the Illumina Human Beadchip HT-12 Arrays.

**Assessment of mitochondrial priming by BH3 profiling**

Xenograft cells permeabilized by digitonin were exposed to BH3 peptides derived from BAD, NOXA, and PUMA proteins, and mitochondrial depolarization measured using the fluorescent dye JC-1, as we have previously described (20). Comparison of mitochondrial depolarization of nonresponders versus responders was performed using a t test (unpaired, 2-tailed).

**In vitro cytotoxicity assays**

The in vitro sensitivity of xenograft cells to ABT-263 was assessed by coculture using hTERT-immortalized primary bone marrow mesenchymal stromal cells (hTERT-MSC), as described previously (29) and detailed in the Supplementary Methods. Briefly, hTERT-MSCs were seeded at 2,000 cells/well in a 384-well plate (Greiner) in serum-free medium (AIM-V; Life Technologies). After 24 hours, 20,000 viable leukemia cells and ABT-263 were added to final concentrations of 1, 2.5, 5, 10, 25, 50, 100 and 1,000 nmol/L in duplicate. After 72 hours of incubation live-cell numbers were determined using 7-AAD analyses by flow cytometry (BD FACSCountII). Data were normalized using SPHERO AccuCount Particles. Examples of flow cytometry analysis are included in Supplementary Fig. S1. To build a predictive model of in vivo ABT-263 sensitivity, we used upper and lower limits of the 95% confidence intervals of the proportion of live cells after exposure to 10 nmol/L of ABT-263 in vitro.

**Statistical analysis**

EFS curves were compared by the log-rank test. Differences in responses to single-agent ABT-263 in vivo between MLL-ALL, BCP-ALL, and T-ALL xenografts were evaluated using one-way ANOVA and a Tukey multiple comparison analysis as well as χ² test. A Pearson correlation test was utilized for all datasets with normal distribution, which included gene expression analysis versus LGD, and MCL1 Illumina mRNA versus MCL1 RT-qPCR mRNA levels. A Spearman correlation test was used to compare MCL1...
RT-qPCR mRNA versus protein expression. Comparison of MCL1 RT-qPCR mRNA and protein levels between non-responders and responders was performed using a Mann–Whitney test. Significance was inferred from tests with \( P \) values lower than 0.05.

Results

Gene expression profiles of ALL xenografts reflect the primary disease

To identify cell and molecular determinants of \textit{in vivo} ABT-263 responses in pediatric ALL, panels of a total of 31 xenografts were established from direct patient explants representative of MLL-ALL (8 infant MLL-ALLs and 1 pediatric MLL-ALL, ALL-3), BCP-ALL (\( n = 7 \)), and T-ALL (\( n = 15 \)) and were characterized by gene expression profiling. The patient demographics focused on high-risk or poor outcome cases: infant MLL-ALL is a known high-risk ALL subtype; the T-ALL panel included 3 xenografts derived from patients with early T-cell precursor (ETP) ALL (ETP-1, ETP-2, and ETP-3), a very high-risk subgroup (30); the BCP-ALL and non-ETP-T-ALL panels included 4 of 7 and 9 of 12 patients, respectively, who had relapsed and/or died from their disease (22, 30, 31). More detailed descriptions of the infant MLL-ALL and expanded T-ALL xenograft panels will be reported elsewhere. Chromosomal translocations in the original biopsy sample, where known, are summarized in Supplementary Table S2.

Unsupervised hierarchical clustering of xenograft basal gene expression profiles revealed 3 distinct branches reflecting each leukemia subtype (Fig. 1A). The MLL-ALL and BCP-ALL panels seemed more closely related and distinct from the T-ALL xenografts. Xenograft ALL-3 was originally classified by immunophenotype as a BCP-ALL but clustered with the T-ALL xenografts. Xenograft ALL-3 was originally classified by immunophenotype as a BCP-ALL but clustered with the T-ALL xenografts. Consequently, we renamed it T-ALL, with 3 MCRs, 11 CRs, and 5 PRs (Table 1 and Fig. 2).

We next identified subtype-specific differentially expressed genes using the LimmaGP (Cowley and colleagues, manuscript in preparation) module in GenePattern, whereby each subtype was compared with the remaining 2 ALL subtypes (1 vs. rest comparison). Subtype-specific genes were determined with a cut-off value of \( FDR < 0.05 \). At this level of stringency, there were 2,141 MLL-ALL specific genes, 643 BCP-ALL specific genes, and 16,692 T-ALL specific genes (Supplementary Table S3). The top 25 differentially expressed probesets between each xenograft panel included previously identified subtype-specific genes, such as MEIS1, ZNF827, and CCNA1 in MLL-ALL (32), MME (CD10) in BCP-ALL, and components of the CD3 receptor (CD3D, CD3E, CD3G, and CD247), CD2 and SH2D1A in T-ALL (ref. 33; Fig. 1B). ALL subtype-specific genesets were identified using GSEA preranked module in GenePattern with \( FDR < 0.05 \) (Supplementary Table S4). For MLL-ALL, the top 4 genesets reflected MLL-specific genesets. For BCP-ALL, 9 B-cell–specific genesets were identified within the top 30 genesets and for T-ALL, 6 T-cell–specific genesets were identified within the top 10 genesets. Therefore, these analyses confirmed the xenograft subtype classification according to the primary disease.

ABT-263 exhibits single-agent \textit{in vivo} efficacy against a broad range of pediatric ALL subtypes

We previously reported the results of \textit{in vivo} ABT-263 testing against a panel of 6 ALL xenografts, with higher sensitivity observed in 2 T-ALL compared with 4 BCP-ALL xenografts (10). To further investigate this possible subtype-specific \textit{in vivo} efficacy of ABT-263 against pediatric ALL, we expanded the analysis to the 31 xenografts described above. ABT-263 significantly delayed the progression of 29 of 31 xenografts tested (Table 1, Fig. 2A–C, and Supplementary Figs. S2–S4). LGDs ranged from 0.5 (ALL-2; \( P = 0.46 \)) to 78 (ALL-31; \( P = 0.0008 \)) days. When stratified according to ALL subtype the median LGDs were 17.9 days for MLL (range 3.1–53.7), 25.8 days for BCP-ALL (range 0.5–37.9), and 29.6 days for T-ALL (range 4.0–78; Fig. 2D and Table 1). There was no significant differential efficacy of ABT-263 against any of the 3 ALL subtypes.

ABT-263 elicited objective responses in 19 of 31 xenografts, with 3 MCRs, 11 CRs, and 5 PRs (Table 1 and Fig. 2). Figure 2E represents the \textit{in vivo} ABT-263 responses of each xenograft panel in a "COMPARE-like" format. In agreement with the LGD data, no significant differences were observed between the ALL subtypes.

We previously showed that the \textit{in vivo} sensitivity of a subset of these xenografts to an induction-type regimen of vincristine, dexamethasone, and \textit{L}-asparaginase (VXL) correlated with the clinical outcome of the patients from whom the xenografts were derived (34). However, the \textit{in vivo} ABT-263 sensitivity of the same subset of xenografts did not correlate with their VXL responses (\( R = 0.46, P = 0.18 \); Supplementary Fig. S5), indicating that ABT-263 can exert significant \textit{in vivo} efficacy against ALL xenografts that are resistant to established drugs.

A complete summary of results is provided in Supplementary Figs. S2 to S4 and Table S5, including total numbers of mice, number of mice that died (or were otherwise excluded), numbers of mice with events and average times to events, LGD values, as well as numbers in each of the ORM categories and "treated over controls" (T/C) values.

\textit{MCL1} gene expression correlates with \textit{in vivo} ABT-263 sensitivity in ALL xenografts

We next analyzed basal BCL2 family gene expression levels in relation to \textit{in vivo} ABT-263 sensitivity across the 31 xenografts using 2 approaches. Both approaches were applied to the entire xenograft cohort as well as all 3 subtypes separately. The first approach assessed the correlation between gene expression and progression delay (LGD; Pearson product moment correlation coefficient), with a positive correlation denoting genes whose higher
expression was associated with ABT-263 sensitivity, and vice versa. Using this approach, MCL1 expression correlated significantly with in vivo ABT-263 sensitivity across the entire xenograft panel ($R = -0.43$, $P = 0.015$; Fig. 3A), indicating that high MCL1 expression was associated with in vivo resistance. BCLXL/XS ($R = 0.79$, $P = 0.01$) and BCLW ($R = 0.69$, $P = 0.039$) showed significant positive correlations between expression and drug sensitivity in the MLL-ALL panel (Supplementary Fig. S6 and Table S6). In contrast, no significant correlations were observed in the BCP-ALL panel (Supplementary Fig. S7 and Table S6) whereas BID levels were negatively correlated with sensitivity in the T-ALL xenografts ($R = -0.52$, $P = 0.046$; Supplementary Fig. S8 and Table S6).

In the second approach, xenografts were stratified according to their ORM into responders (PRs, CRs, and MCRs) and nonresponders (PDs and SDs) and differentially expressed genes were identified using an unadjusted $P$ value of $\leq 0.05$ from an empirical Bayes moderated $t$ test (28). MCL1 was the BCL2 family member with the strongest differential expression among the entire xenograft panel between responders and nonresponders (Fig. 3B and Supplementary Table S6). This was also the case for the MLL-ALL ($P = 0.027$) and BCP-ALL ($P = 0.008$) subtypes (Supplementary Table S6).
Figs. S6 and S7 and Table S6). Among the MLL-ALL xenografts, differential expression of BIM ($P = 0.044$); also reached statistical significance, being paradoxically lower in the responders compared with nonresponders (Supplementary Fig. S6 and Table S6). Similarly, HRK ($P = 0.005$) was significantly increased in the BCP-ALL responders (Supplementary Fig. S7 and Table S6), whereas BAK ($P = 0.03$) and NOXA ($P = 0.049$) were significantly increased in the T-ALL nonresponders and responders, respectively (Supplementary Fig. S8 and Table S6).

Both of the above analysis approaches had also been applied to the entire 34,694 genes represented on the Illumina Beadchip HT-12 arrays, although no genes satisfied the significance or FDR cutoff criteria (data not shown).

Because MCL1 expression was the strongest overall predictor of in vivo ABT-263 response across the entire panel of 31 xenografts, we next assessed MCL1 expression at the mRNA and protein levels. Although RT-qPCR analysis showed a significantly higher MCL1 expression in the nonresponders (Fig. 3C), this difference was not confirmed by increased MCL1 protein levels (Fig. 3D and Supplementary Fig. S9). Despite no significant difference in MCL1 protein levels between non-responders and responders, MCL1 mRNA levels correlated between qRT-PCR and microarray (Fig. 3E), and MCL1 protein levels significantly correlated with MCL1 mRNA expression measured by qRT-PCR (Fig. 3F). MCL1 protein expression was also investigated after exposure of 2 non-responders and 2 responders to ABT-263 in vitro, however no consistent differences were observed (data not shown). Because of these discrepancies in MCL1 protein expression, we next assessed BCL2 family protein

### Table 1. In vivo responses of pediatric ALL xenografts to ABT-263

<table>
<thead>
<tr>
<th>ALL lineage</th>
<th>Xenograft ID</th>
<th>Vehicle control</th>
<th>ABT-263</th>
<th>LGD (days)</th>
<th>$P$ value (log-rank)</th>
<th>Median ORM</th>
<th>ORM heatmap</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLL-ALL</td>
<td>MLL-1</td>
<td>19.2</td>
<td>39.7</td>
<td>20.5</td>
<td>0.0003</td>
<td>6</td>
<td>PR</td>
</tr>
<tr>
<td></td>
<td>MLL-2</td>
<td>14.8</td>
<td>26.8</td>
<td>12.0</td>
<td>0.002</td>
<td>6</td>
<td>PR</td>
</tr>
<tr>
<td></td>
<td>MLL-3</td>
<td>8.3</td>
<td>34.1</td>
<td>25.8</td>
<td>0.0002</td>
<td>8</td>
<td>CR</td>
</tr>
<tr>
<td></td>
<td>MLL-5</td>
<td>8.7</td>
<td>14.6</td>
<td>5.9</td>
<td>0.0006</td>
<td>2</td>
<td>PD2</td>
</tr>
<tr>
<td></td>
<td>MLL-6</td>
<td>6.3</td>
<td>29.1</td>
<td>22.8</td>
<td>0.0002</td>
<td>2</td>
<td>PD2</td>
</tr>
<tr>
<td></td>
<td>MLL-7</td>
<td>12.7</td>
<td>30.6</td>
<td>17.9</td>
<td>0.0002</td>
<td>6</td>
<td>PR</td>
</tr>
<tr>
<td></td>
<td>MLL-8</td>
<td>13.7</td>
<td>16.8</td>
<td>3.1</td>
<td>0.0002</td>
<td>0</td>
<td>PD1</td>
</tr>
<tr>
<td></td>
<td>MLL-14</td>
<td>8.7</td>
<td>24.9</td>
<td>16.2</td>
<td>0.0006</td>
<td>2</td>
<td>PD2</td>
</tr>
<tr>
<td></td>
<td>ALL-3</td>
<td>9.6</td>
<td>63.3</td>
<td>53.7</td>
<td>0.008</td>
<td>8</td>
<td>CR</td>
</tr>
<tr>
<td>BCP-ALL</td>
<td>ALL-2$^a$</td>
<td>4.7</td>
<td>5.2</td>
<td>0.5</td>
<td>0.46</td>
<td>0</td>
<td>PD1</td>
</tr>
<tr>
<td></td>
<td>ALL-4$^a$</td>
<td>1.5</td>
<td>28.6</td>
<td>27.1</td>
<td>0.029</td>
<td>2</td>
<td>PD2</td>
</tr>
<tr>
<td></td>
<td>ALL-7</td>
<td>6.8</td>
<td>31.4</td>
<td>24.6</td>
<td>0.0003</td>
<td>8</td>
<td>CR</td>
</tr>
<tr>
<td></td>
<td>ALL-10</td>
<td>7.5</td>
<td>33.3</td>
<td>25.8</td>
<td>0.008</td>
<td>8</td>
<td>CR</td>
</tr>
<tr>
<td></td>
<td>ALL-11</td>
<td>20.2</td>
<td>56.3</td>
<td>36.1</td>
<td>0.008</td>
<td>8</td>
<td>CR</td>
</tr>
<tr>
<td></td>
<td>ALL-17$^a$</td>
<td>2.5</td>
<td>24.2</td>
<td>21.7</td>
<td>0.0001</td>
<td>2</td>
<td>PD2</td>
</tr>
<tr>
<td></td>
<td>ALL-19$^a$</td>
<td>5.0</td>
<td>&gt;43</td>
<td>&gt;37.9</td>
<td>0.002</td>
<td>8</td>
<td>CR</td>
</tr>
<tr>
<td>T-ALL</td>
<td>ALL-8$^a$</td>
<td>11.8</td>
<td>62.8</td>
<td>51.0</td>
<td>0.004</td>
<td>10</td>
<td>MCR</td>
</tr>
<tr>
<td></td>
<td>ALL-16$^a$</td>
<td>18.2</td>
<td>85.5</td>
<td>67.3</td>
<td>0.0002</td>
<td>10</td>
<td>MCR</td>
</tr>
<tr>
<td></td>
<td>ALL-27</td>
<td>3.3</td>
<td>19.6</td>
<td>16.3</td>
<td>0.30</td>
<td>2</td>
<td>PD2</td>
</tr>
<tr>
<td></td>
<td>ALL-29</td>
<td>4.4</td>
<td>48.8</td>
<td>44.0</td>
<td>0.0002</td>
<td>8</td>
<td>CR</td>
</tr>
<tr>
<td></td>
<td>ALL-30</td>
<td>5.3</td>
<td>16.2</td>
<td>10.9</td>
<td>0.004</td>
<td>2</td>
<td>PD2</td>
</tr>
<tr>
<td></td>
<td>ALL-31</td>
<td>13.9</td>
<td>91.9</td>
<td>78.0</td>
<td>0.0008</td>
<td>10</td>
<td>MCR</td>
</tr>
<tr>
<td></td>
<td>ALL-32$^b$</td>
<td>6.7</td>
<td>18.9</td>
<td>12.2</td>
<td>0.0002</td>
<td>2</td>
<td>PD2</td>
</tr>
<tr>
<td></td>
<td>ALL-33$^b$</td>
<td>6.0</td>
<td>37.8</td>
<td>31.8</td>
<td>0.0003</td>
<td>8</td>
<td>CR</td>
</tr>
<tr>
<td></td>
<td>ALL-42</td>
<td>1.9</td>
<td>5.9</td>
<td>4.0</td>
<td>0.0002</td>
<td>2</td>
<td>PD2</td>
</tr>
<tr>
<td></td>
<td>ALL-43</td>
<td>14.0</td>
<td>43.9</td>
<td>29.9</td>
<td>0.0002</td>
<td>8</td>
<td>CR</td>
</tr>
<tr>
<td></td>
<td>ALL-44</td>
<td>11.4</td>
<td>27.2</td>
<td>15.8</td>
<td>0.0005</td>
<td>2</td>
<td>PD2</td>
</tr>
<tr>
<td></td>
<td>ALL-47</td>
<td>11.9</td>
<td>41.5</td>
<td>29.6</td>
<td>&lt;0.0001</td>
<td>8</td>
<td>CR</td>
</tr>
<tr>
<td>ETP-ALL</td>
<td>ETP-1</td>
<td>7.8</td>
<td>23.8</td>
<td>16.0</td>
<td>0.0002</td>
<td>6</td>
<td>PR</td>
</tr>
<tr>
<td></td>
<td>ETP-2</td>
<td>14.2</td>
<td>44.7</td>
<td>30.5</td>
<td>0.0002</td>
<td>8</td>
<td>CR</td>
</tr>
<tr>
<td></td>
<td>ETP-3</td>
<td>20.0</td>
<td>32.5</td>
<td>12.5</td>
<td>0.0002</td>
<td>6</td>
<td>PR</td>
</tr>
</tbody>
</table>

$^a$In vivo ABT-263 sensitivity data previously reported (10).

$^b$Dose of ABT-263 reduced to 75 mg/kg on day 9 (ALL-32) or day 13 (ALL-33) because of toxicity.
function with respect to \textit{in vivo} ABT-263 sensitivity using BH3 profiling of the entire xenograft panel.

**BH3 profiling identifies MCL1 function as a determinant of \textit{in vivo} ABT-263 sensitivity**

The mitochondrial priming assay measures mitochondrial sensitivity to peptides derived from the BH3 domains of proapoptotic BCL2 family proteins. The \textit{in vivo} ABT-263 responses of the xenografts were then compared with the status of mitochondrial priming by the ability of BH3 peptides derived from PUMA, BAD, and NOXA to cause mitochondrial depolarization in xenograft cells. PUMA BH3 interacts promiscuously with all 5 antiapoptotic proteins, BAD BH3 interacts with BCL2, BCLXL, and BCLW (like ABT-263), whereas NOXA BH3 interacts only with MCL1. Out of the 3 peptides, only NOXA-induced mitochondrial depolarization significantly discriminated between non-responder and responder groups (Fig. 3G–I), implicating MCL1 protein function as a major determinant of \textit{in vivo} ABT-263 response. The extent of mitochondrial depolarization did not correlate with leukemia progression (LGD) for any of the peptides tested (data not shown).

**\textit{In vitro} ABT-263 sensitivity of ALL xenografts predicts their \textit{in vivo} responses**

We next tested whether the \textit{in vitro} ABT-263 responses of ALL xenografts predicted their \textit{in vivo} sensitivity. Using a coculture method (29), a predictive model was built using a training subset of 17 xenografts (Fig. 4A). Exposure of xenograft cells to a range of ABT-263 concentrations (1 μmol/L–1 nmol/L) revealed that 10 nmol/L gave the best discrimination between \textit{in vivo} non-responders and responders (Fig. 4B and Supplementary Fig. S10). Using the 95%
Figure 3. Cell and molecular determinants of the in vivo sensitivity of pediatric ALL xenografts to ABT-263. A, xenografts (columns) were ordered by increasing LGD from left to right, with each row representing a BCL2 family member. B, xenografts were stratified into nonresponders (NR) and responders (R) then ordered by increasing LGD from left to right within each category. The colors in the heatmaps represent the relative expression per gene across all samples. Red indicates relative high expression and blue indicates relative low expression. C, comparison of MCL1 mRNA expression between NR and R by RT-qPCR. D, comparison of MCL1 protein expression between NR and R by immunoblot. E, correlation between MCL1 mRNA expression by microarray and RT-qPCR. F, correlation between MCL1 protein expression by immunoblot and MCL1 mRNA expression by RT-qPCR. In C–F, each data point represents a single xenograft. G–I, the percentage of mitochondrial depolarization induced by BH3 peptides derived from (G) PUMA, (H) BAD, and (I) NOXA peptides in xenograft cells was plotted for individual xenografts stratified as NR or R according to Table 1.

The in vivo ABT-263 responses of the remaining 11 xenografts, constituting the test set, were initially blinded. When the test set was classified using the predictive model 7 xenografts were correctly classified according to their in vivo

The confidence intervals of both responder and nonresponder groups, a three-tier classification was created. It was established that >14.9% live cells after drug treatment (lower limit of confidence interval of nonresponders) stratified xenografts as nonresponders, <8.0% (upper limit of confidence interval of responders) as responders, and those in between were considered as unclassified. Three xenografts (MLL-1, MLL-7, and ETP-3) were excluded from the analysis because their survival was not supported by the coculture assay.

www.aacrjournals.org Clin Cancer Res; 20(17) September 1, 2014 4527

Published OnlineFirst July 10, 2014; DOI: 10.1158/1078-0432.CCR-14-0259
ABT-263 responses (Fig. 4C), 2 were incorrectly classified and 2 were unclassified. The sensitivity and specificity of the predictive method (Supplementary Fig. S11) were 50% and 100%, respectively (leaving out of this assessment the 2 unclassified xenografts).

Live and dead cell analysis at 72 hours for each xenograft is shown in Supplementary Fig. S12, which includes the % live/dead cells, the absolute numbers of live/dead cells, and the number of live cells seeded and harvested. In addition, a comparison of the coculture system versus single cell suspension was performed with 5 xenograft samples, and revealed the importance of the coculture assay to assess ALL xenograft cell sensitivity to ABT-263 in vitro (Supplementary Fig. S13).

**In vivo efficacy of ABT-263 in combination with established chemotherapeutic drugs**

We next sought to test the efficacy of ABT-263 in paired combinations with established drugs against xenografts representative of the 2 most common pediatric ALL subtypes, BCP-ALL and T-ALL. Three xenografts were selected based on their range of single-agent ABT-263 responses (ALL-2, resistant; ALL-19, intermediate; ALL-31, sensitive; Table 1). Preliminary tolerability experiments showed that it was necessary to attenuate the dose of ABT-263 to 25 mg/kg when combined with vincristine, and to 50 mg/kg when combined with dexamethasone and L-asparaginase. Of the 3 xenografts tested, the combination of ABT-263 with vincristine caused therapeutic enhancement only in ALL-31 and although for ALL-19 it induced an LGD of >73.7 days, because variation among individual mice it did not reach statistical significance versus ABT-263 alone (Supplementary Tables S7 and S8 and Fig. S14). The dexamethasone/ABT-263 combination did not exert therapeutic enhancement for any xenografts, whereas ABT-263 in combination with L-asparaginase exerted therapeutic enhancement in ALL-31, and the P value approached significance for ALL-19. Thus, the expectation that ABT-263 would broadly enhance the in vivo efficacy of established chemotherapeutic drugs was not met.

**Discussion**

We report the utilization of a large panel of ALL xenografts to define the cell and molecular determinants of in vivo ABT-263 responses using gene expression profiling, BH3 profiling, and in vitro coculture cytotoxicity assays. The principal findings of this study are: (i) ABT-263 is effective in vivo as a single agent against pediatric ALL xenografts; (ii) MCL1
Determinants of In Vivo ABT-263 Response

gene expression and MCL1 protein function correlate with in vivo ABT-263 sensitivity; and (iii) an in vitro coculture cytotoxicity assay is able to predict in vivo ABT-263 responses of ALL xenografts with a high level of sensitivity and specificity.

Xenograft models of pediatric ALL are recognized to accurately recapitulate several cellular and molecular features of the original disease, including blast morphology, immunophenotype, clonal selection, gene expression profiles, and genetic lesions (22, 31, 35–37). Our results describe the subtype classification of a large panel of xenografts, which were appropriately clustered into MLL-ALL, BCP-, and T-ALL subtypes by gene expression profiling. Differentially expressed genes and GSEA analysis within each subtype were consistent with the primary disease state, and a BCP-ALL xenograft previously established from a teenage female (31) was reclassified as an MLL-ALL based in this analysis. Subclusters within each xenograft subtype were also consistent with specific chromosomal translocations. For example, all 4 of the MLL-ALL xenografts harboring a translocation involving chromosome 19 coclustered in 1 subbranch of MLL-ALL, whereas the 3 ETP-ALLs appeared under 1 subbranch of T-ALL.

Despite our previous report of ABT-263 efficacy testing against 6 ALL xenografts indicating a preferential effect against 2 T-ALL xenografts (10), in this study ABT-263 exhibited a broad spectrum of in vivo efficacy with no apparent subtype specificity, and induced regressions in 19 of 31 xenografts. However, all 3 xenografts that achieved MCRs were T-ALL, suggesting that ABT-263 may be particularly useful for the treatment of aggressive T-ALL cases. ABT-263 also induced regressions (2 PRs and 1 CR) in the 3 ETP-ALL xenografts. ETP-ALL arises from a subset of thymocytes that are recent emigrants from the bone marrow to the thymus, retaining stem cell–like features and multilineage differential potential, and is a particularly aggressive and refractory T-ALL subtype (30).

In this study, we also attempted to identify cell and molecular signatures that could be used to predict in vivo responsiveness to single-agent ABT-263. Studies primarily carried out using cultured cell lines have consistently identified elevated MCL1 expression to be associated with resistance to ABT-737 and ABT-263 (14–16, 38). Although our study is no exception, we believe this to be the first report to strongly implicate MCL1 in resistance to ABT-263 in vivo using a large panel of direct-patient explants established as continuous xenografts. This relationship only became apparent when we restricted the gene expression analysis to the BCL2 gene family, a finding that we attribute to the heterogeneity in gene expression profiles between and within each xenograft subtype. Nevertheless, the molecular determinants of in vivo ABT-263 sensitivity in pediatric ALL are complex since additional, and occasionally paradoxically, BCL2 family genes were significantly associated with ABT-263 responses within individual xenograft subtypes.

A confounding factor in our analysis was that, although MCL1 mRNA expression determined by RT-qPCR correlated with MCL1 protein expression and microarray data, MCL1 protein expression did not reach statistical significance between the ABT-263 responders and non-responders despite a higher trend in the nonresponders (Fig. 3D). Technical issues associated with the harvesting, purifying, and cryopreservation of spleen-derived cells and the very short half-life of MCL1 (<1 hour; refs. 39 and 40) may have contributed to this lack of correlation because of MCL1 protein degradation (Supplementary Fig. S15). Moreover, Gao and Koide reported that there are 2 MCL1 splice variants regulated by SF3B1, one with proapoptotic and another with antiapoptotic functions (41). However, we found no correlation between SF3B1 levels and drug sensitivity (data not shown). Similarly, Boiani and colleagues showed that the HSP70 protein BAG3 stabilizes MCL1, thereby extending its half-life (42). Similarly, we found no correlation between BAG3 expression and in vivo ABT-263 efficacy (data not shown). Because of the aforementioned complexity associated with correlating gene expression profiles and MCL1 protein expression with in vivo ABT-263 sensitivity, we next tested the established functional readouts of mitochondrial priming status by BH3 profiling and in vivo chemosensitivity testing. Mitochondrial depolarization induced by NOXA peptide correlated with in vivo ABT-263 sensitivity (Fig. 3I), thereby strongly implicating MCL1 function. However, this finding differs from a previous report that identified a correlation between BIM, but not NOXA, peptide-induced mitochondrial depolarization and clinical complete response to conventional therapy using pediatric ALL biopsy samples (20).

Murine or human bone marrow–derived stromal cells cocultured with MSC-hTERT cells provided a superior survival of pediatric ALL cells (29, 43) whereas both coculture and tetrazolium dye-based assays have been used for chemosensitivity testing in this disease (44–46). In our study, ABT-263 sensitivity of pediatric ALL xenograft cells cocultured with MSC-hTERT cells provided a sensitive (50%) and highly specific (100%) prediction of in vivo response. This model accurately predicted resistance, because all 5 of the in vivo nonresponders were correctly identified. However, although both of the xenografts predicted to be responders were correct, the model incorrectly predicted no response in 2 of the in vivo responders. Therefore, this model could be further refined, because although it might accurately predict which patients are unlikely to respond, it can potentially fail to identify a subset of patients who may benefit from such treatment. Nevertheless, we believe this to be the first report of a functional assay that is able to accurately predict in vivo single-agent ABT-263 responses.

Despite substantial evidence, primarily using cultured cell lines, that ABT-737 and ABT-263 can potentiate the effects of standard chemotherapeutic drugs both in vitro and in vivo (9, 16, 47, 48), using stringent criteria we only observed Therapeutic Enhancement in 2 instances in which ABT-263 was combined with 3 established drugs against 3 xenografts. Although our results are not sufficient to make broad conclusions for combining ABT-263 with these established drugs for patient management, we reason that this divergence from previous reports is because of the necessity
to attenuate the ABT-263 dose in all of the combinations (down to 25 mg/kg in the case of vincristine), while maintaining the maximal ABT-263 dose (100 mg/kg) in the single-agent arms. Future investigations in which small-molecule BCL2 inhibitors with reduced thrombocytopenic effects, such as ABT-199 (49, 50), are combined with established drugs in pediatric ALL may prove more beneficial.

In summary, BCL2-targeted agents appear as a promising class of anticancer drugs for the treatment of pediatric ALL, with no apparent specificity across MLL-ALL, BCP-ALL, or T-ALL subtypes. MCL1 expression and function seem to be important determinants of in vivo ABT-263 sensitivity, although an in vitro coculture assay predicted in vivo ABT-263 responses with high sensitivity and specificity. This combined cell and molecular analysis provides a proof-of-concept approach for prioritizing other novel drugs for pediatric ALL clinical trials, and for the identification of biomarkers predictive of in vivo response.

Disclosure of Potential Conflicts of Interest
Anthony Letai is a consultant/advisory board member for AbbVie. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions
Conception and design: H. Carol, C. Sarmah, P.J. Houghton, M.A. Smith, R.B. Lock
Acquisition of data (provided animals, acutated analyses, acutated and managed patients, provided facilities, etc.): S. Suryani, H. Carol, T.N. Chonghaile, V. Frismantas, L. High, B. Szymanska, K. Evans, I. Boehm, E. Tonna, L. Jones, D.M. Manesh, A. Letai, J.-P. Bourquin
Writing, review, and/or revision of the manuscript: S. Suryani, H. Carol, C. Sarmah, B. Bornhauser, M.J. Cowley, B. Szymanska, K. Evans, E. Tonna, L. Jones, P.J. Houghton, M.A. Smith, R.B. Lock
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): S. Suryani, K. Evans, E. Tonna, R.T. Kurnasheva
Study supervision: S. Suryani, H. Carol, R.B. Lock

Acknowledgments
The authors thank AbbVie Inc. for providing ABT-263, the Tissue Resources Core Facility of St Jude Children’s Research Hospital (Memphis, TN) for the provision of primary ETP-ALL samples, and Dr Dario Campana (St Jude Children’s Research Hospital, Memphis, TN) for providing KIF4A-immortalized MSCs. Children’s Cancer Institute Australia for Medical Research is affiliated with the University of New South Wales and The Sydney Children’s Hospitals Network.

Grant Support
This research was funded by grants from the National Cancer Institute (NOI-CM-42216 and NOI-CM-91001-03); the Leukaemia Foundation of Australia the Cancer Council New South Wales; the Krebsliga Zurich to B. Bornhauser, and the Foundation Kind und Krebs, the Hanne Liebermann Stiftung, and the Swiss National Research Foundation to J.-P. Bourquin. S. Suryani is supported by Postdoctoral Fellowships from the Leukaemia Foundation of Australia and the Cure Cancer Australia Foundation, and an Early Career Fellowship from the Cancer Institute NSW. R.B. Lock is supported by a Fellowship from the National Health and Medical Research Council.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 31, 2014; revised June 10, 2014; accepted June 29, 2014; published OnlineFirst July 10, 2014.

References

Cincancerres.aacrjournals.org on April 15, 2017. © 2014 American Association for Cancer Research.
Clinical Cancer Research

Cell and Molecular Determinants of In Vivo Efficacy of the BH3 Mimetic ABT-263 against Pediatric Acute Lymphoblastic Leukemia Xenografts

Santi Suryani, Hernan Carol, Triona Ni Chonghaile, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-14-0259

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2014/07/18/1078-0432.CCR-14-0259.DC1

Cited articles
This article cites 49 articles, 24 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/20/17/4520.full.html#ref-list-1

Citing articles
This article has been cited by 8 HighWire-hosted articles. Access the articles at:
/content/20/17/4520.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.