Clinical significance of P-glycoprotein expression and function for response to induction chemotherapy, relapse rate and overall survival in acute leukemia

CHRISTIAN WUCHTER,* LEONID KARANAEW,* VELIA RUPPERT,† MARTIN SCHRAPPE,* THOMAS BÜCHNER,* CLAUDIA SCHÖCH,† TORSTEN HAFERLACH,† JOCHEN HARBOTT,† RICHARD RATEI,* BERND DÖRKEN,* WOLF-DIETER LUDWIG*  
*Dept. of Hematology, Oncology, and Tumor Immunology, Robert-Rössle-Clinic, Charité, Humboldt-University of Berlin; †Dept. of Pediatrics, MH Hannover; ‡Dept. of Internal Medicine A, University of Münster; @Dept. of Internal Medicine III, Ludwig Maximilians-University of München; ^Oncogenetic Laboratory, Children’s Hospital, Justus-Liebig-University of Gießen, Germany

ABSTRACT

Background and Objectives. A multidrug-resistance (MDR) phenotype mediated by P-glycoprotein (P-gp) contributes to chemotherapy failure in acute leukemia. However, the exact prognostic significance of this resistance mechanism is still unclear, mostly due to methodologic problems in P-gp detection. We therefore investigated whether P-gp expression levels or functional P-gp activity better predict response to induction chemotherapy, relapse rate and overall survival in acute leukemia.

Design and Methods. We examined cell samples of 121 adults with de novo acute myeloid leukemia (AML) and 102 children with newly diagnosed acute lymphoblastic leukemia (ALL) for P-gp expression and functional P-gp activity by flow cytometry. P-gp expression was determined by the rhodamine 123 (rh123)-efflux test (AML n=121, ALL n=102) and P-gp expression levels using the P-gp specific monoclonal antibodies (moabs) MRK-16 (AML n=51, ALL n=31), 4.E3 (AML n=35, ALL n=32), or UIC-2 (AML n=68, ALL n=50). We correlated our findings with the immunophenotype, FAB morphology, cytogenetics and clinical data of the examined patients.

Results. P-gp expression levels as detected by MRK-16 and 4.E3 were very low and did not differ between AML and ALL as estimated using relative fluorescence intensity (RFI) values and D-values by Kolmogorov-Smirnov (KS) statistics. For moab UIC-2, P-gp expression levels were higher in AML than in ALL. Within AML, moab UIC-2 mainly reacted with myelomonocytic-differentiated leukemic cells of the FAB M4/5 subtypes. No correlation between P-gp expression levels as detected by MRK-16, 4.E3 or UIC-2 and the response to induction chemotherapy or relapse rate, both in AML and ALL, was observed. However, a prognostic impact of P-gp expression levels on overall survival in AML was seen for moab MRK-16. Moreover, within AML, P-gp function was higher in immature blast cells as defined by immunophenotype and FAB morphology and correlated with response to induction chemotherapy, relapse rate, overall survival as well as cytogenetic risk groups. In ALL, the overall functional P-gp activity was lower than in AML and did not correlate with immunophenotypic subgroups, response to induction chemotherapy, relapse rate or overall survival.

Interpretation and Conclusions. Our data demonstrate a prognostic impact of P-gp in AML but not ALL and indicate that the functional rh123-efflux assay should be preferred for flow cytometric P-gp evaluation in acute leukemia compared with P-gp expression analysis by monoclonal antibodies.

Key words: acute leukemia, multidrug-resistance, P-glycoprotein expression and function

A multidrug resistance (MDR) phenotype, mainly mediated by the efflux pump P-glycoprotein (P-gp), contributes to chemotherapy failure in acute leukemia. However, despite numerous studies within the last ten years, the exact expression pattern and the prognostic value of P-gp in acute leukemia are not clear so far. Several studies found an association between P-gp expression and/or P-gp function and a CD34 positive immature phenotype in acute myeloid leukemia (AML).1-3 However, other studies did not confirm these data.6,7 Whereas several investigators found a correlation between high P-gp expression levels and/or P-gp function and a poor response to chemotherapy in AML1-3,8-10 and acute lymphoblastic leukemia (ALL),11-13 other studies revealed no clear prognostic impact for P-gp in acute leukemia.14-16

One of the main reasons for these contradictory results may be methodologic problems in P-gp detection. A workshop of leading MDR experts concluded that the accurate measurement of low levels of P-gp expression remains an elusive goal, and that it is still uncertain which level of P-gp detection correlates with the acquisition of drug resistance by a patient’s tumor cells.17 In an editorial to a round-table debate on MDR it was mentioned that serious faulty methodology is the major cause of rejection of papers dealing with MDR.18 For MDR analyses in acute leukemia, the P-gp detection methods currently most often used are i) measurement of P-gp function by
efflux studies, ii) P-gp expression levels by monoclonal antibodies (moabs) and iii) MDR-1 gene expression (encoding for P-gp) by reverse-transcriptase polymerase chain-reaction (RT-PCR).17 However, each of these methods has specific disadvantages. P-gp expression analysis may be hampered by non-specific staining and weak P-gp expression levels. Functional MDR-tests may be not specific for P-gp activity, and RT-PCR as a very sensitive method may detect low, clinically irrelevant MDR-1 mRNA expression levels. Moreover, neither mRNA nor protein expression detection methods allow an estimation of functional P-gp activity. Therefore, there have been proposals to combine different P-gp detection methods.18,19 Single-cell detection methods (e.g. P-gp expression, P-gp activity by flow cytometry) were recommended as they may afford more reliability and accuracy than bulk methods (e.g. RT-PCR) in that the former are generally able to discriminate tumor from normal cells.19

To determine whether P-gp expression levels or functional P-gp activity better predict response to induction chemotherapy, relapse rate and overall survival in acute leukemia, we examined a large series of leukemic cell samples by flow cytometry for P-gp function, as detected by the rhodamine 123 (rh123)-efflux test, and P-gp expression levels using the P-gp specific moabs MRK-16, 4.3E or UIC-2. As reference immunophenotyping laboratory of the German multicenter AM LC86/92 and ALL-BFM21-23 trials, we included adults with de novo myeloid leukemia (Table 1A) and children with newly diagnosed acute lymphoblastic leukemia (Table 1B) in this study. To determine the value of P-gp detection methods for high-risk groups in childhood ALL, we took particular care to include cell samples from patients with poor response to initial prednisone therapy and/or noresponse to initial induction chemotherapy in this study (Table 1B). We correlated our findings with maturation stage (immunophenotype, FAB morphology) and cytogenetic features of the leukemic cells examined. Moreover, the prognostic value of P-gp expression and P-gp function for response to induction chemotherapy, relapse rate and overall survival was determined.

**Design and Methods**

**Patients and cell samples**

We examined either freshly obtained or cryopreserved bone marrow or peripheral blood samples from 121 adults with de novo AML and 102 children with newly diagnosed ALL. No differences between freshly obtained and prior cryopreserved cell samples were seen for the investigated parameters. Diagnosis of AML and ALL was made morphologically according to the FAB criteria24,25 and immunophenotypically according to the EGIIL recommendations.26 All samples contained more than 80% leukemic cells based on morphologic and immunologic criteria. Of 106 adult AML patients (included in the German multicenter AM LC86/92 trials26), initial laboratory parameters (WBC, platelet count, lactate dehydrogenase) and the response to daunorubicin and cytosine arabinoside-based induction chemotherapy were available for correlation with the experimental data (Table 1A). Of 73 AML patients with complete remission (CR) after induction chemotherapy, 27 are in continuous complete remission (CCR) (mean event-free-survival, EFS: 20 months), and 41 patients have relapsed (mean EFS: 12 months). No follow-up data from 5 patients with CR after induction chemotherapy were available for further analysis. Of 92 children with ALL (included in the multicenter ALL-BFM-90/95 trials21-23), the response to initial prednisone therapy (prednisone response) and induction chemotherapy as well as initial laboratory parameters (WBC, platelet count, hemoglobin) were available for correlation with the experimental data (Table 1B). Of 80 ALL patients with CR after induction chemotherapy, 53 patients (41 B-lineage ALL, 12 T-ALL) are in CCR (mean EFS: 57

<table>
<thead>
<tr>
<th>Number of patients</th>
<th>106 (55 males, 51 females)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, mean (range)</td>
<td>54 (22-78)</td>
</tr>
<tr>
<td>Initial WBC x10^9/L, mean (range)</td>
<td>64.2 (1.5-450) (n=105)</td>
</tr>
<tr>
<td>Initial platelet count x10^9/L, mean (range)</td>
<td>82 (8-387) (n=104)</td>
</tr>
<tr>
<td>Initial LDH units/L, mean (range)</td>
<td>776 (161-3678) (n=86)</td>
</tr>
<tr>
<td>FAB morphology M0 (n=6), M1 (n=20), M2 (n=37), M3 (n=3), M4 (n=24), M5 (n=15)</td>
<td></td>
</tr>
<tr>
<td>Cytogenetic risk groups good (n=4), intermediate (n=33), poor (n=9)</td>
<td></td>
</tr>
<tr>
<td>Response to induction chemotherapy early death (n=10) complete remission (n=73) non-response (n=23)</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1B. Characteristics of 92 examined children with newly diagnosed ALL treated according to the ALL-BFM 90/95 protocol.**

<table>
<thead>
<tr>
<th>Number of patients</th>
<th>92 (55 males, 37 females)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, median (range)</td>
<td>7 (0-18)</td>
</tr>
<tr>
<td>Initial WBC x10^9/L, mean (range)</td>
<td>146.9 (1.4-1200)</td>
</tr>
<tr>
<td>Initial platelet count x10^9/L, mean (range)</td>
<td>71 (7-404)</td>
</tr>
<tr>
<td>Initial hemoglobin g/dL, mean (range)</td>
<td>9.3 (4.6-17.7)</td>
</tr>
<tr>
<td>Immunophenotype B-lineage ALL (n=47), T-lineage ALL (n=25)</td>
<td></td>
</tr>
<tr>
<td>Response to initial prednisone therapy good response (n=69; 47 B-lineage ALL, 12 T-lineage ALL) poor response (n=33; 20 B-lineage ALL, 13 T-lineage ALL)</td>
<td></td>
</tr>
<tr>
<td>Response to induction chemotherapy complete remission (n=80; 61 B-lineage ALL, 19 T-lineage ALL) non-response (n=12; 6 B-lineage ALL, 6 T-lineage ALL)</td>
<td></td>
</tr>
</tbody>
</table>
months), 20 (14 B-lineage ALL, 6 T-ALL) have relapsed (mean EFS: 17 months). No follow-up data from 7 patients with CR were available for further analysis.

As a positive control for P-gp expression and P-gp function analysis, the MDR-1 gene transfected subline HT29mdr1 of the colon carcinoma cell line HT29 was kindly provided by Dr. Gottesman (National Cancer Institute, NIH, Bethesda, MD, USA).

Immunophenotyping
Leukemic cells from heparinized bone marrow or peripheral blood samples were isolated by Ficoll-Hypaque (Pharmacia, Biotech, Freiburg, Germany) density gradient centrifugation, and leukemia-associated cell-surface antigens were detected by a panel of mAbs either by direct or indirect immunofluorescence assay as previously described.28,29 Cell samples were analyzed by flow cytometry (FACScan; Becton Dickinson, San Diego, USA) using the Cell-Quest software program (Becton Dickinson). Cell samples were considered positive for a specific antigen if the antigen was expressed on at least 20% of the leukemic cells (20-cut-off level).28,30

P-gp expression by flow cytometry
In the first part of our study, leukemic cells were stained with the unconjugated mAbs MRK-1631 (5 µg/10⁶ cells; IgG2a; Dianova, Hamburg, Germany) and 4.E3 (32) (2.5 µg/10⁶ cells; IgG2a; Signet, Dedham, USA) by indirect immunofluorescence using FITC-labeled goat-anti-mouse antibody (Dianova) as second layer. Thirty-five AML cell samples and 31 ALL cell samples were analyzed for both mAbs. Sixteen AML cell samples were analyzed for mAb MRK-16 alone, and one ALL cell sample was analyzed for mAb 4.E3 alone. Because of the weak P-gp expression levels as detected by these mAbs, we used the PE-conjugated mAb UIC223 (0.5 µg/10⁶ cells; IgG1; Dianova) for the second part of our study. In this part, 68 AML cell samples and 50 ALL cell samples were analyzed for P-gp expression using mAb UIC-2. All three mAbs examined detect extracellular epitopes of the MDR1-P-gp and do not crossreact with MDR3-P-gp.34 At least 10,000 cells per sample were acquired and analyzed by flow cytometry as described above. Expression of antigens was quantified as relative fluorescence intensity (RFI), determined by the ratio of median fluorescence intensity of cells specifically stained for either MRK-16, 4.E3, or UIC-2 to median fluorescence intensity of cells stained with appropriate isotype-matched control antibodies.4 P-gp expression was also evaluated using the Kolmogorow-Smirnov (KS) test.35,36 As described in previous studies, D-values ≥0.15 obtained with the KS test were considered positive.35,36 Non-viable cells in each sample were excluded from analysis by propidium iodide staining (0.3 µg/mL; Sigma, Deisenhofen, Germany). Non-specific Fc-receptor-mediated binding of mAbs was blocked by preincubation of the leukemic cells with a polyclonal rabbit serum (Gibco BRL, Paisley, UK).

Detection of P-glycoprotein function
P-gp function was investigated using the rh123-efflux test as described before.3 Briefly, cells were stained with rh123 (0.5 µg/mL) for 30 minutes at 37°C. After washing twice, cells were incubated either with or without verapamil (5 µM) for two hours in RPMI medium at 37°C and then analyzed by flow cytometry. Non-viable cells were excluded from analysis by propidium iodide staining. Cell samples were considered P-gp positive if more than 10% of the cells stained with rh123 alone revealed a lower rh123-fluorescence intensity as compared with cells in the corresponding control sample with verapamil.1,37

Cytogenetics
Chromosome analyses were performed on metaphases from short term (24h, 48h) cultures of pre-treatment bone marrow and/or peripheral blood cells. Cell cultivation and chromosome preparation were carried out according to standard protocols. G-banding was used and the chromosomes interpreted according to ISCN (1995) nomenclature.38 AML patients were grouped into three different risk categories: good: t(8;21), inv(16)/t(16;16), t(15;17); intermediate: normal karyotype, other abnormalities; poor: -5/-5q-, -7/-7q-, 11q23-abnormalities, inv(3)/t(3;3), t(9;22), t(6;9), 17p-abnormalities, karyotypes with three or more numerical or structural abnormalities. Cytogenetic data of all patients were grouped into cases with hyperdiploidy, normal karyotype or other abnormalities.

Statistical analysis
Differences in the expression levels of P-gp as detected by the mAbs MRK-16, 4.E3 and UIC-2 between the immunophenotypic subgroups (e.g. CD34pos. vs. CD34neg.) and FAB subtypes (M0/1 vs. M2-M5) were evaluated with the Mann-Whitney U-test or chi-square test (Pearson’s coefficient). All statistical correlations with FAB morphology and immunophenotypic data as well as clinical data were made for each mAb separately. Differences in the rh123-efflux between the immunophenotypic subgroups (e.g. CD34pos. vs. CD34neg.) and FAB subtypes (M0/1 vs. M2-M5) were evaluated using the chi-square test (Pearson’s coefficient). P-gp expression levels were correlated with the extent of rh123-efflux in individual cell samples by bivariate correlation (Spearman’s correlation coefficient rs). Univariate analyses of the experimental findings with cytogenetic groups, response to initial prednisone therapy, induction chemotherapy and relapse status were performed using either the Mann-Whitney U-test (P-gp expression levels) or chi-square test (P-gp function). Differences in the overall survival of P-gp positive and negative AML or ALL subgroups were evaluated using the log rank test. All statistical analyses were done with the SPSS software program.

Results
Constitutive P-gp expression levels in acute leukemia
Eighty-two cell samples were analyzed for P-gp expression using the mAb MRK-16 as described in
the Design and Methods section. P-gp surface expression was generally weak and ranged between 1.0 and 1.9 RFI (mean RFI 1.2±0.2). Sixty-seven cell samples were analyzed for P-gp expression using the moab 4.E3, again as described in the Design and Methods section. P-gp surface expression in this case too was generally weak and ranged between 1.0 and 2.0 RFI (mean RFI 1.1±0.2). No differences in P-gp expression levels as detected by the moabs MRK-16 and 4.E3 between AML and ALL as well as B- and T-lineage ALL were observed (Table 2). There was no correlation between MRK-16 and 4.E3 reactivity and maturation stage, either in AML (Table 3) or in ALL (data not shown).

In individual cell samples, P-gp expression levels, as detected by MRK-16 or 4.E3, were significantly correlated (rs =0.7, n=61, p<0.001). Reactivity of MRK-16 and 4.E3 was also tested with the MDR-1 gene transfected cell line HT29mdr1. P-gp expression of HT29mdr1 cells was high (RFI=10.4) compared with that in fresh leukemic cells (Figure 1). Additionally, 118 cell samples were analyzed for P-gp expression using the moab UIC-2. P-gp surface expression ranged between 1.0 and 9.6 RFI (mean RFI 2.0±1.7). P-gp expression levels were significantly higher in AML than in ALL (Table 2). In AML, UIC-2 reacted more often with myelomonocytic-differentiated AML of the FAB M4/5 subtypes than with FAB M0-3 subtypes (M4/5: n=24, 3.9 ±2.6 RFI vs. M0-3: n=39, 1.8 ±0.8 RFI, p<0.0009). There was no correlation between UIC-2 reactivity and maturation stage in ALL (data not shown).

We also analyzed P-gp expression as a dichotomous variable using different RFI threshold levels for positivity (e.g. ≥1.2, ≥2.0). However, for neither moab MRK-16 nor moab 4.E3 did we find any significant correlation with maturation stage of the leukemic cells nor any differences between AML and

| Table 2. P-gp expression levels and P-gp function in acute leukemia: correlation with immunological subtypes (ALL) and FAB morphology (AML) [P-gp expression: mean RFI±SD; P-gp function: number of samples with >10% rh123 efflux positive cells/number of samples tested (% of samples)]. |
|---|---|---|---|---|---|---|
| | P-gp expression by monoclonal antibodies | | | P-gp function by rh123-efflux | | |
| | MRK-16 | n | p-value | 4.E3 | n | p-value | UIC-2 | n | p-value | p-value |
| AL total | 1.2±0.2 | 82 | ns | 1.1±0.2 | 67 | ns | 2.0±1.6 | 118 | ns | 58/233 (26%) |
| AML | 1.2±0.2 | 51 | ns | 1.1±0.1 | 35 | ns | 2.5±1.9 | 68 | ns | 48/121 (40%) |
| ALL | 1.1±0.2 | 31 | ns | 1.1±0.2 | 32 | ns | 1.4±0.8 | 50 | ns | 10/102 (10%) |
| B-lineage ALL | 1.1±0.2 | 23 | ns | 1.1±0.2 | 24 | ns | 1.4±0.9 | 41 | ns | 8/73 (11%) |
| T-lineage ALL | 1.2±0.2 | 8 | 1.1±0.2 | 8 | ns | 1.3±0.6 | 9 | ns | 2/29 (7%) |
| AML FAB total | 1.2±0.2 | 44 | ns | 1.1±0.1 | 32 | ns | 2.6±2.0 | 63 | ns | 44/112 (39%) |
| AML FAB M0/1 | 1.2±0.2 | 15 | ns | 1.1±0.2 | 13 | ns | 1.8±0.7 | 13 | ns | 19/30 (63%) |
| AML FAB M2-5 | 1.2±0.2 | 29 | ns | 1.1±0.1 | 19 | ns | 2.8±2.1 | 50 | ns | 25/82 (30%) |

Abbreviation: ns, not significant.

| Table 3. P-gp expression levels and P-gp function in adult de novo acute myeloid leukemia: correlation with immunophenotype [P-gp expression: mean RFI±SD; P-gp function: number of samples with >10% rh123 efflux positive cells/number of samples tested (% of samples)]. |
|---|---|---|---|---|---|---|
| | P-gp expression by monoclonal antibodies | | | P-gp function by rh123-efflux | | |
| | MRK-16 | n | p-value | 4.E3 | n | p-value | UIC-2 | n | p-value | p-value |
| CD34 pos. | 1.2±0.2 | 25 | ns | 1.1±0.1 | 19 | rs | 2.0±0.8 | 26 | rs | 39/52 (75%) |
| CD34 neg. | 1.2±0.2 | 26 | ns | 1.1±0.1 | 16 | rs | 2.7±2.4 | 42 | ns | 9/69 (13%) |
| TdT pos. | 1.2±0.2 | 11 | ns | 1.1±0.0 | 8 | rs | 2.4±1.3 | 6 | ns | 14/18 (78%) |
| TdT neg. | 1.2±0.2 | 35 | ns | 1.1±0.1 | 24 | ns | 2.5±2.0 | 58 | ns | 31/96 (32%) |
| CD7 pos. | 1.3±0.3 | 12 | ns | 1.1±0.1 | 7 | ns | 1.9±0.9 | 7 | ns | 14/20 (70%) |
| CD7 neg. | 1.1±0.2 | 36 | ns | 1.1±0.1 | 26 | ns | 2.5±2.5 | 61 | ns | 13/100 (33%) |
| CD19 pos. | 1.2±0.3 | 45 | ns | 1.1±0.0 | 1 | ns | 1.9±0.5 | 3 | ns | 5/6 (83%) |
| CD19 neg. | 1.2±0.3 | 11 | ns | 1.1±0.0 | 13 | ns | 2.5±2.5 | 65 | ns | 42/114 (37%) |
| CD15 pos. | 1.2±0.2 | 12 | ns | 1.1±0.1 | 17 | ns | 2.8±2.4 | 34 | ns | 13/27 (49%) |
| CD15 neg. | 1.2±0.2 | 35 | ns | 1.1±0.1 | 26 | ns | 2.1±2.3 | 33 | ns | 32/68 (47%) |
| CD65 pos. | 1.1±0.1 | 25 | ns | 1.1±0.1 | 18 | ns | 2.7±2.2 | 45 | ns | 22/74 (30%) |
| CD65 neg. | 1.2±0.2 | 22 | ns | 1.1±0.1 | 15 | ns | 2.0±1.1 | 23 | ns | 24/44 (55%) |
| CD14 pos. | 1.2±0.1 | 8 | ns | 1.1±0.1 | 5 | ns | 3.0±2.5 | 14 | ns | 4/23 (17%) |
| CD14 neg. | 1.2±0.1 | 37 | ns | 1.1±0.1 | 26 | ns | 2.3±1.8 | 54 | ns | 40/93 (43%) |
| CD64 pos. | 1.2±0.2 | 20 | ns | 1.1±0.1 | 12 | ns | 1.8±2.0 | 41 | ns | 16/61 (26%) |
| CD64 neg. | 1.2±0.2 | 17 | ns | 1.1±0.2 | 13 | ns | 2.1±3.8 | 25 | ns | 25/43 (58%) |

Abbreviations: ns, not significant; nd, not done.
P-glycoprotein expression and function in acute leukemia

ALL and the immunophenotypic subtypes in ALL (data not shown). Using a RFI threshold level of ≥ 2.0, moab UIC-2 reacted more often with AML than ALL cell samples (27 (40%) /68 vs. 6 (12%) /50; p<0.001). In AML, myelomonocytic-differentiated AML of the FAB M4/5 subtypes were more often positive for moab UIC-2 compared with FAB M0-3 subtypes (17 (71%) /24 vs. 10 (26%) /39; p<0.0005).

P-gp expression was also analyzed using the Kolmogorow-Smirnov (KS) test. However, neither by using the obtained D-values as continuous variables nor by defining a D-value ≥ 0.15 for positivity, did we find any significant correlation between the moabs MRK-16 or 4.E3 and the maturation stage of the leukemic cells, AML and ALL or the immunophenotypic subtypes in ALL (data not shown). However, the D-values obtained with moab UIC-2 were significantly higher in AML than in ALL (mean D-value: 0.38±0.22 (n=68) vs. 0.13±0.14 (n=50); p<0.0001).

In AML, the D-values obtained with moab UIC-2 were significantly higher in myelomonocytic-differentiated AML of the FAB M4/5 subtypes than in the FAB M0-3 subtypes (0.50±0.23 (n=24) vs. 0.33±0.18 (n=39); p<0.002).

Correlation of P-gp expression levels with functional P-gp activity

The only significant correlation between rh123-efflux (% of rh123-efflux positive cells) and P-gp expression levels in individual cell samples was seen for rh123-efflux and P-gp expression levels as detected by moab MRK-16 in AML using RFI values for P-gp expression analysis (rs= 0.3, (n=51), p<0.05/D-values: rs= 0.3, (n=51), p=0.06), but not in ALL. P-gp expression levels as detected by moabs 4.E3 or UIC-2 were not significantly correlated with rh123-efflux in individual cell samples, either in AML or in ALL (data not shown). P-gp expression levels as detected by moabs MRK-16 and 4.E3 correlated in individual cell samples using D-values (r=0.6, (n=66), p<0.001) as well as using RFI values (r=0.4, (n=66), p<0.001).

Correlation of the experimental findings with cytogenetic data of the examined patients

Cytogenetic data of 58 adults with de novo AMI were available for correlation with the investigated parameters. Cytogenetic risk groups were defined as stated in the Design and Methods section. Leukemic cells of patients in the poor cytogenetic risk group had a significantly higher P-gp function than leukemic cells of patients in the good/intermediate cytogenetic risk groups (Table 4). No clear-cut differences in the P-gp

Figure 1. P-gp expression levels in a representative AML cell sample and the MDR-1 gene transfected cell line HT29 detected by the moabs MRK-16 and 4.E3 as described in the Design and Methods section.

Figure 2. P-gp function in de novo acute leukemia

P-gp function was analyzed in 223 cell samples as described in the Design and Methods section. Figure 2 shows representative examples of rh123-efflux in acute leukemia cell samples. The range of rh123-efflux positive cells/sample was between 0 and 90% (mean 11±13%). Fifty-eight of 223 (26%) cell samples revealed more than 10% rh123-efflux positive cells/sample. P-gp function was significantly higher in AML than in ALL (p<0.000001) (Table 2). No differences in P-gp function between B- and T-lineage ALL were observed (Table 2). P-glycoprotein function correlated with maturation stage in de novo AMI: immature FAB M0/1 subtypes revealed a higher P-gp function than more mature FAB M2-5 subtypes (p<0.002) (Table 2). Functional P-gp activity in AMI was associated with an immature immunophenotype characterized by significantly higher CD34 expression (p<0.00001), coexpression of the lymphoid markers TdT (p<0.0003), CD7 (p<0.002), CD19 (p<0.03) and lower expression of the myelomonocytic differentiation markers CD15 (p<0.03), CD65s (p<0.008), CD14 (p<0.03) and CD64 (p<0.002) (Table 3). There was no correlation between P-gp function and maturation stage in ALL (data not shown).

Correlation of P-gp expression levels with functional P-gp activity

The only significant correlation between rh123-efflux (% of rh123-efflux positive cells) and P-gp expression levels in individual cell samples was seen for rh123-efflux and P-gp expression levels as detected by moab MRK-16 in AML using RFI values for P-gp expression analysis (rs= 0.3, (n=51), p<0.05/D-values: rs= 0.3, (n=51), p=0.06), but not in ALL. P-gp expression levels as detected by moabs 4.E3 or UIC-2 were not significantly correlated with rh123-efflux in individual cell samples, either in AML or in ALL (data not shown). P-gp expression levels as detected by moabs MRK-16 and 4.E3 correlated in individual cell samples using D-values (r=0.6, (n=66), p<0.001) as well as using RFI values (r=0.4, (n=66), p<0.001).
expression levels as detected by the moabs MRK-16, 4.E3 or UIC-2, were found between the cytogenetic risk groups in AML, either using RFI values (Table 4) or D-values (data not shown).

Cytogenetic data of 23 children with ALL were available for correlation with the investigated parameters. Two of 7 cell samples with hyperdiploidy were rh123-efflux positive, 7 cell samples with normal karyotype and 8 cell samples with other cytogenetic anomalies [e.g., t(4;11), del(6q)] were all rh123-efflux negative (not significant). No differences in the P-gp expression levels, as detected by the moabs MRK-16, 4.E3 or UIC-2, were found between the cytogenetic groups in ALL (data not shown).

Correlation of the experimental findings with clinical data of the examined patients

In adults, correlation of complete remission after induction chemotherapy was correlated with a significantly lower P-gp function than that in non-responding patients (Table 5A). There was no correlation between P-gp expression levels and response to induction chemotherapy using either RFI values (Table 5A) or D-values as obtained by the KS test (data not shown). Seventeen (41%) of the 41 patients who relapsed after complete remission were rh123-efflux positive, whereas only 5 (19%) of 27 in CCR were rh123-efflux positive (p<0.05). P-gp expression levels, as detected by moab UIC-2, were slightly lower in relapsed patients than in those in CCR. However, this difference was not statistically significant (2.5±0.8 RFI (n=30) vs. 2.8±0.5 RFI (n=21); p=0.31).

Neither for moab MRK-16 (11 relapses, 5 CCR) nor moab 4.E3 (9 relapses, 3 CCR), were differences in P-gp expression levels between these two groups seen (data not shown). Overall survival (OS) was significantly longer in patients with rh123-efflux negative AML cell samples (mean OS: 34 months) compared with patients with rh123-efflux positive cell samples (mean OS: 11 months) (Figure 3). There was no significant correlation between P-gp expression levels, as detected by the moabs 4.E3 or UIC-2, and overall survival (data not shown). However, using the KS test and a D-value of ≥1.15 as the cut-off level for P-gp positivity, moab MRK-16 positive AML cases (n=10; mean OS: 7 months) revealed a significant lower OS than moab MRK-16 negative AML cases (n=18; mean OS: 33 months) (log rank test = 0.0227). No correlation between the experimental data and initial laboratory parameters in AML (WBC, platelet count, LDH) was found (data not shown).

In childhood ALL, neither P-gp function nor P-gp expression as detected by moab UIC-2 was significantly correlated with response to initial prednisone therapy (Table 5B) or response to induction chemotherapy (Table 5C). None of the cell samples from 20 patients with relapse after complete remission was rh123-efflux positive, whereas 6 (11%) of 53 cell samples from patients in CCR were rh123-efflux positive (difference not significant). P-gp expression levels, as detected by moab UIC-2, were higher in patients with relapse than in those in CCR. However, this difference was not statistically significant (RFI: 1.6±0.8 (n=11) vs. 1.2±0.2 (n=26); p=0.31/D-values: 0.17±0.20 (n=11) vs. 0.10±0.08...
Overall survival (OS) was longer in patients with rh123-efflux positive ALL cell samples compared with those with rh123-efflux negative cell samples, however, this difference was not statistically significant (Figure 3). There was no significant correlation between P-gp expression levels, as detected by the moab UIC-2, and overall survival (data not shown). No statistical correlation between expression levels and clinical data was performed for moabs MRK-16 and 4.E3 for childhood ALL patients because of the small number of patients examined in each of the clinical groups.

**Discussion**

In this study we investigated P-gp expression as well as its functional activity in 223 cell samples from children and adults with de novo acute leukemia. We correlated these parameters with immunophenotypic features, FAB morphology, cytogenetics and clinical data of the examined patients.

In accordance with other studies, we observed a correlation between high functional P-gp activity and CD34 positivity,3-5 immature FAB subtypes,39 and poor cytogenetic risk features40 in AML. In normal hematopoiesis, CD34 positive stem cells express functional P-gp. P-gp may protect these pluripotent progenitor cells from toxic substances or may inhibit differentiation induced by cytokines by effluxing these substances out of the cell.41 However, the exact role of P-gp in hematopoietic stem cells is not yet clear. During myeloid differentiation, P-gp expression is downregulated as mature monocytes and granulocytes are P-gp negative.42 Based on our findings, we conclude that the maturation-dependent differences in functional P-gp activity of immature and more mature AML cells resemble the maturation-dependent decrease in P-gp function of normal myeloid progenitor cells. P-gp downregulation must also be postulated during lymphoid differentiation as most thymocytes are negative for P-gp.43 However, mature T- and B-cell subsets as well as NK-cells...
again express P-gp. The functional role of P-gp in these cells is no longer clearly defined. P-gp may be involved in the cellular transport of cytokines. In accordance with another study, we found only low functional P-gp activity in ALL cells compared with AML cells. A higher P-gp expression in T-lineage compared with B-lineage ALL has been observed. However, we found no clear differences, either in P-gp expression or P-gp function, between these two ALL subgroups (Table 2).

The prognostic significance of P-gp function for response to induction chemotherapy, relapse frequency and event-free survival as well as overall survival in our series of adult de novo AML patients is consistent with the results of several studies on the influence of P-gp on therapy response and prognosis in AML. We have recently shown that functional P-gp activity is more predictive of response to induction chemotherapy in adult de novo AML than the expression levels of the apoptosis-regulating molecules CD95 (Fas/APO-1), Bcl-2 and Bax.

We found no correlation between P-gp expression/function and response to induction chemotherapy, relapse rate, as well as event-free survival and overall survival in childhood ALL. This contrasts partly with results of a recently published prospective study showing P-gp expression, as detected by moabs C219 and 4E3 by APAAP-staining, to be an independent prognostic factor for prediction of relapse but not response to induction chemotherapy in childhood ALL. In accordance with Dhooge et al., we also found higher P-gp expression levels in adult samples of patients with relapse compared with samples from patients in CR. However, this difference was not statistically significant. Despite the fact that we especially examined cell samples of high-risk patients (Table 1B), the functional P-gp activity was low in childhood ALL and did not differ significantly between cell samples from patients with relapse and samples from patients in CR. Moreover, we found no correlation between P-gp expression or P-gp function and overall survival in childhood ALL (Figure 3). These findings support the suggestion of den Boer et al. that in childhood ALL other resistance mechanisms may be more important for chemotherapy failure than P-gp mediated chemoresistance. On the other hand, the low P-gp activity observed may be one of the reasons for the good response to chemotherapy treatment in childhood ALL.

In accordance with other studies, we found only weak P-gp expression levels in primary acute leukemia cell samples compared with drug-resistant cell lines. In a large French multicenter MDR-trial, the generally weak P-gp expression levels (as detected by the moabs M RK-16, 4E3 and UIC-2) were only comparable between the centers if they used the more sensitive RFI values instead of the percentage of positive cells for flow cytometric P-gp expression analysis. Because of the weak P-gp expression levels, other investigators also preferred the RFI values to percentage values for P-gp expression analysis. For the same reason, den Boer et al. explicitly recommended not to use moab 4E3 for P-gp detection by flow cytometry. The results of our study support this recommendation. The weak P-gp expression levels in primary acute leukemia cells that we detected by M RK-16 and 4E3 were most likely not due to methodologic failures, as we detected high P-gp expression levels in the M DR-1 gene-transfected subline of the cell line HT29 (Figure 1). It was shown that values of P-gp expression detected depend on the antibody concentrations used. This fact at least partly explains discrepancies between reported P-gp expression levels in acute leukemia cell samples detected by the same P-gp specific moab in different laboratories and studies.

The KS test may be a useful tool for P-gp expression analysis, as it is an often used statistical approach for the analysis of weak P-gp expression levels in acute leukemia. Interestingly, the D-values obtained by the KS test but not RFI values revealed a significant correlation between P-gp expression levels as detected by moab M RK-16 and overall survival in AML in our study (Figure 3). A higher fluorescence intensity may be achieved by conjugation of the P-gp moabs with the fluorochrome phycoerythrin (PE) instead of with FITC. However, our results with the directly PE-conjugated moab UIC-2 show that PE-conjugation could lead to non-specific binding to membrane cytocytic differentiated AML cells. Intracellular P-gp expression analysis by moab UIC-2 may improve the specificity of this antibody as a recent study revealed a significant correlation between intracellular P-gp expression levels as detected by moab UIC-2 and functional P-gp activity as well as clinical data in AML.

Neither MDR1-mRNA detection by RT-PCR nor detection of P-gp expression levels allows clear estimation of the functional activity of the P-gp molecule. Post-translational modifications and regulations may influence the functional potential of P-gp. In accordance with the findings of Leith et al., we did not find a clear correlation between P-gp expression and P-gp function as detected by the rh123-efflux test in individual cell samples. This lack of correlation may be explained by the fact that the rh123-efflux test is more sensitive in the detection of low levels of drug resistance than P-gp expression analysis. Moreover, other proteins such as MRP (MDR related protein) and LRP (lung resistance protein) contribute to the MDR-phenotype of acute leukemia cells. MRP and P-gp both belong to the ATP-binding cassette superfamily of membrane-bound transporter molecules. Thus, the rh123-efflux test may detect, besides P-gp function, MRP-activity. In a recent paper by Leith et al., P-gp but not MRP or LRP expression levels were correlated with rh123-efflux in individual cell samples of adults with AML. These authors mentioned that the rh123-efflux of MRP positive cells was slower than that of P-gp positive cells. Measurement of rh123-efflux after 90-120 minutes should, therefore, be a good marker of functional P-gp activity. Whether measurement of rh123-efflux after 3 hours correlates with functional MRP activity is not yet clear. The newly developed calcim-AM test may be a better functional test, with prognostic significance, for distinguishing P-gp and MRP activity within a single leukemic cell sample. To our knowledge, no functional test for LRP detection has yet been described.
Our results show that the rh123-efflux test, using the 10% cut-off level for flow-cytometric analysis, is an effective tool for adequate detection of P-gp function in acute leukemia. This fact should facilitate the integration of functional P-gp assessment in clinically oriented immunophenotyping. Moreover, P-gp function analysis by rh123-efflux test is at least ten times cheaper than detection of P-gp expression levels by P-gp specific mAbs. An eventual improvement of P-gp expression analysis by using higher concentrations of the P-gp mAbs/sample would lead to a further increase in costs. Our results, in accordance with other studies, show that the rh123-efflux test can be performed on freshly obtained samples and cryopreserved cell samples with similar results. However, we found that early apoptotic cells show a verapamil-insensitive rh123-efflux (unpublished observation). Therefore, all rh123-efflux tests must include control samples with P-gp inhibitor substances such as verapamil or cyclosporin A. Moreover, non-viable cells should be excluded from analysis by propidium iodide staining.

We conclude that neither P-gp expression level nor P-gp function correlate with response to induction chemotherapy, relapse rate or overall survival in childhood ALL. In contrast, functional P-gp activity gives a better prediction of response to induction chemotherapy, relapse, and overall survival in adult de novo AML than P-gp expression levels do. We therefore recommend including the functional rh123-efflux test as a cheap and accurate flow cytometric method for P-gp evaluation in clinically oriented AML immunophenotyping.

Funding
This work was supported by grants from the "Deutsche José Carreras Leukämie Stiftung" (IJCLS 1998/NAT-3 to CW) and the "Deutsche Leukämie-Forschungshilfe" (DLFH-98.04 to VR).

Contributions and Acknowledgments
CW, LK and WDL were responsible for the design of this study and for the interpretation of the experimental results. CW, LK and VR participated in the data acquisition and statistical analysis of the experimental findings. CS performed the cytogenetic studies in AML and ALL cases. All the authors were involved in revision of the manuscript and approved the final version to be submitted. We would like to thank K. Ganzel, G. Czerwony, M. Martin and K. Liebezeit for their excellent technical assistance. The cell samples included in this study were sent from various hospitals in Germany participating in the ongoing ALL-BFM (coordinators: M. Schnapp and H. Riehm, H annoer) and AML-CG (coordinator: T. Büchner, M üsster) trials. We thank all clinicians providing cell samples for our investigations.

Disclosures
Conflict of interest: none.
Redundant publications: we previously investigated in adult AML the expression of apoptosis-regulating molecules (CD95, Bcl-2, Bax) in the context of P-glycoprotein function and focused on their prognostic significance for response to induction chemotherapy (Wuchter et al., Leukemia 1999, 13:1943). In contrast, the present article focuses on P-glycoprotein detection methods (P-glycoprotein expression vs. P-glycoprotein function) and their prognostic value (response to induction chemotherapy, relapse rate, overall survival) in both AML and ALL.

Manuscript processing
Manuscript received November 12, 1999; accepted May 11, 2000.

Potential implications for clinical practice
- In adult de novo AML, P-gp activity should be considered as an important prognostic factor for response to chemotherapeutic treatment. This study underlines the necessity for prospective P-gp evaluation with AML therapy trials.
- In childhood ALL, functional P-gp activity is low and does not correlate with clinical outcome. Other resistance mechanisms seem to be more important for chemotherapy failure in childhood ALL.
- For flow cytometric P-gp evaluation in acute leukemia, the functional rhodamine 123-efflux test for P-gp activity analysis should be preferred to P-gp expression analysis by monoclonal antibodies.

References


30. Chaudhary PM, Roninson IB. Expression and activity of P-glycoprotein, a multidrug efflux pump, in human