Prognostic Correlation of Plasma Cell Acid Phosphatase and β-Glucuronidase in Multiple Myeloma: A Southwest Oncology Group Study

By Sheikh M. Saeed, Donna Stock-Novack, Rita Pohlod, John Crowley, and Sydney E. Salmon

In 1982 a randomized trial of either alternating or syncopated VMCP/VBAP regimens for the treatment of active multiple myeloma was begun (Southwest Oncology Group Study 8229/30). A concurrent investigation was undertaken to evaluate the clinical importance and significance of cytochemically stainable plasma cell acid phosphatase (AP) and β-glucuronidase enzymes (BG). Pretreatment bone marrow aspirates were available for analysis from 399 patients for AP and 398 patients for BG. The AP scores ranged between 42 and 395, and the BG scores ranged between 1 and 346. There was a significant increase of AP \( (P = .001) \) and BG \( (P = .002) \) in multiple myeloma as compared with a set of patients with benign plasmacytosis. The enzyme scores did not significantly relate to Ig idotype of myeloma or other prognostic variables except that the BG scores varied significantly with the level of albumin \( (P = .03) \) and hemoglobin \( (P = .01) \). Analysis of patient groups with different levels of enzyme scores showed that 61 of 398 patients with an AP score of less than 130 had a poorer median survival of 1.7 versus 2.8 years for patients with higher scores \( (P = .001) \). In the multivariate analysis of survival, low AP score was an important prognostic factor \( (P = .006) \), but BG did not contribute significantly. It is suggested that the subset of patients presenting with low AP should be considered for specialized or more aggressive therapy.

© 1991 by The American Society of Hematology.

LYSOSOMES are important enzyme-containing organelles of hematopoietic cells such as granulocytes, lymphocytes, and plasma cells. It has been reported that acid phosphatase (AP), β-glucuronidase (BG), and nonspecific esterase enzymes are increased in myeloma cells compared with normal plasma cells and plasma cells in monoclonal gamopathies of unknown significance (MGUS). On the other hand, ATP-ase activity has been found to be reduced in myeloma cells. It has also been suggested that the levels and the activity of such osteoclastic and hydrolytic enzymes as AP and BG may be involved in normal bone resorption and in the production of bone lesions in multiple myeloma (MM). In 1980, Bataille et al reported that for 38 patients studied, AP activity correlated with disease activity in myeloma. They also found that the average AP score decreased significantly in remission as compared with pretreatment levels. In contrast, Hoffman et al, in a longitudinal study of five treated myeloma patients, reported AP scores to be very stable and independent of the disease activity. With respect to BG, Seigneurin et al reported for 34 myeloma patients that BG was significantly increased in MM and that there was correlation between BG score and Durie-Salmon myeloma stage. In 1982, a new combined modality protocol for therapy of MM patients was initiated in the Southwest Oncology Group (SWOG 8229/30). As a component of the protocol, a prospective study of myeloma cell AP and BG enzymes was included to evaluate the potential clinical significance and prognostic implications of these enzymes as evaluated cytochemically. This report summarizes laboratory findings of the study and suggests that low plasma cell AP defines a group of myeloma patients who have a significantly poor prognosis.

MATERIALS AND METHODS

Between 1982 and 1987, 621 previously untreated patients with active MM were entered into a randomized trial of two arms of alternating and syncopated VMCP/VBAP regimens (SWOG study 8229/30). Approval was obtained from the Institutional Review Boards for these studies. Informed consent was provided according to the Declaration of Helsinki. Six hundred fourteen patients were eligible for the overall study and the clinical results of the study were recently reported. Studies conducted at the central laboratory at Henry Ford Hospital established that both AP and BG in the air-dried smears of myeloma marrows remain stable without fixation or staining for 2 weeks at room temperature. Thereafter, the enzyme content begins to decrease and the values become unreliable by the third week. To be eligible for entry of patients into this enzyme study, SWOG participants submitted air-dried, unfixed bone marrow (BM) aspirate smears to the central laboratory within 1 week of BM aspiration. These smears had to show adequate overall cellularity and plasma cell content to be able to confirm a diagnosis of MM. To establish a baseline and to compare enzyme activity in plasmacytosis (including MGUS) versus MM, extra BM smears from selected Henry Ford Hospital patients were processed and studied in a manner similar to the myeloma cases. The patients with plasmacytosis had 3% to 8% plasma cells in normocellular or hypercellular BM and there was no clinical or laboratory evidence of underlying myeloma, lymphoma, or amyloidosis. Staining of BM smears from patients with plasmacytosis was performed for AP and BG in 34 (8 MGUS) and 37 (10 MGUS) cases, respectively.

Because of the strict requirements for sample quality and time constraints on sample analysis, 399 patient samples were accepted for pretreatment AP scoring and 398 for BG scoring. As evidenced by comparable demographics and survival, there was no selection bias between the group of patients with enzyme values versus the ones without enzyme values.

From the Henry Ford Hospital, Detroit, MI; the Southwest Oncology Group Statistical Center, Seattle, WA; and the University of Arizona Cancer Center, Tucson.

Submitted November 30, 1990; accepted August 16, 1991.

Supported in part by the following Public Health Services Cooperative Agreement grants awarded by the National Cancer Institute, DHHS: CA-04915, CA-37429, CA-13612, CA-28862, CA-13238, CA-04919, CA-04915, CA-16385, CA-35261, CA-03899, CA-35900, CA-22411, CA-04920, CA-35128, CA-12123, CA-35117, CA-35431, CA-36020, CA-12644, CA-46113, CA-14028, CA-37445, CA-35274, CA-27057, CA-35438, CA-35995, CA-35176, CA-35438, CA-32734, CA-30998, CA-46136, CA-35178, CA-35996, CA-35262, CA-35119, CA-35158, CA-35084, CA-35429, CA-35192, CA-32102.

Address reprint requests to the Southwest Oncology Group (SWOG 8229/30), Operations Office, 5430 Fredericksburg Rd, Suite 618, San Antonio, TX, 78229-6197.

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

© 1991 by The American Society of Hematology.
Cytochemical Staining

Cytochemical staining was performed on air-dried BM aspirate smears. AP was shown by using Naphthol AS-BI phosphate as the substrate and Fast Garnet GBC salt as the coupler. BG was shown using Naphthol AS-BI β-D-glucuronic acid as substrate and hexazotized pararosaniline as the coupling agent.

Enzyme Scoring Method

Three years before the start of this protocol study, the enzyme scoring method of Cassuto et al. with and without the modifications recommended by Hoffman et al. had been tried in our laboratory for scoring both AP and BG enzymes. We had found that the methods were cumbersome and the scores were difficult to reproduce by different technologists. A simpler and more reproducible scoring method was sought. Because most investigators are familiar and adept at using Kaplow’s method of scoring leukocyte alkaline phosphatase (LAP), we adopted a method that is similar to LAP scoring. For the scoring of AP or BG, 100 consecutive plasma cells were microscopically examined and scores were assigned based on the amount of plasma cell cytoplasm occupied by enzyme granules. A plasma cell with five or fewer small granules and colorless cytoplasm was scored as 0. Cells with cytoplasmic granules occupying less than 25%, 25% to 49%, 50% to 75%, or greater than 75% of the cytoplasm were assigned scores of 1+, 2+, 3+, or 4+, respectively (Fig 1). The scores of 100 consecutive cells were added for an aggregate value in each patient sample. The first 50 samples were scored independently by Pohlod and Saeed and the scores were reproducible within ±7%. The scoring of all subsequent samples was performed by Pohlod and reviewed by Saeed.

Statistical Methods

General methods. Analyses were performed on an IBM 4381 mainframe computer using version 5.16 of SAS software. Fluctuations in the sample size were due to missing data. As the number of variables examined together at a particular time was increased, the number of records with missing values also increased.

All patients received identical induction chemotherapy drugs administered with either of two schedules. There was no difference in response or survival by induction chemotherapy (P > .73), therefore, data on all patients were pooled for cytochemical evaluation.

Relationship of enzyme scores with other prognostic variables. Previous analysis of this data base had shown that significant predictors of survival were serum β-2 microglobulin (SB2M), albumin, calcium, age, and serum creatinine. Other investigators have found relationships between AP and/or BG with MM light chain type, bone lesions, stage of disease, and hemoglobin, but not with type of M-protein and hemoglobin. Therefore, we evaluated the univariate relationship between AP and BG and these pretreatment variables. Other pretreatment variables examined in this data analysis included performance status, sex, and race. Differences in enzyme scores between patient groups were tested with the Kruskal-Wallis statistic calculated by the SAS procedure PROC NPARIWAY. Scores of myeloma patients were compared with scores of patients with benign plasmacytosis by discriminant analysis.

Relationship of enzyme scores with response and survival. The method of Kaplan and Meier was used to compute the survival curves and to estimate median survival. Survival was measured from the date the patient was registered on the study. Differences in survival curves were evaluated with the log-rank test. The relationship of enzyme scores with response were assessed using the χ² test.

Multivariate analysis. Response was analyzed using logistic regression (LOGIST procedure) while survival was modeled using the Cox proportional hazards model (PHGLM procedure). In both of these procedures, candidate prognostic variables were entered into the model in a step-up manner, where entry required a maximum adjusted P value of .05 and retention required a maximum adjusted P value of .05 once the variable was in the model.

RESULTS

Myeloma Versus Benign Plasmacytosis

The distribution of the AP and BG enzyme scores in plasmacytosis, MGUS, and MM is shown in Table 1. The range of scores for AP or BG in MM was wider than in
Table 1. Enzyme Scores in Plasmacytosis, MGUS, and MM

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasmacytosis</td>
<td>26</td>
<td>153</td>
<td>61</td>
<td>143</td>
<td>39-280</td>
</tr>
<tr>
<td>MGUS</td>
<td>8</td>
<td>156</td>
<td>63</td>
<td>160</td>
<td>73-227</td>
</tr>
<tr>
<td>MM</td>
<td>396</td>
<td>189</td>
<td>55</td>
<td>186</td>
<td>42-395</td>
</tr>
<tr>
<td>BG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasmacytosis</td>
<td>27</td>
<td>91</td>
<td>41</td>
<td>95</td>
<td>27-196</td>
</tr>
<tr>
<td>MGUS</td>
<td>10</td>
<td>109</td>
<td>33</td>
<td>114</td>
<td>52-164</td>
</tr>
<tr>
<td>MM</td>
<td>394</td>
<td>125</td>
<td>42</td>
<td>113</td>
<td>1-346</td>
</tr>
</tbody>
</table>

Pretreatment Findings in MM Patients

Acid phosphatase. There was no difference in AP scores between patients with IgG, IgA, and Bence Jones myeloma nor was there a significant difference between patients grouped by κ versus λ light chain.

Tests for differences in AP based on creatinine and bone lesions suggest possible differences, whereas similar tests based on serum SB2M, albumin, calcium, age, hemoglobin, and clinical stage did not show a particular trend. Patients with low (≤2.0 mg/dL) creatinine had a mean AP of 192, while those with high (≥2.0 mg/dL) creatinine had a mean of 181 (P < .08). The 313 patients with some bone involvement had a slightly higher mean AP score of 193 as compared with 175 for the 63 patients without bone involvement (P < .08).

β-Glucuronidase. Tests for differences in BG based on SB2M, albumin, and hemoglobin were suggestive of differences. Patients with low SB2M (<6 μg/mL) had a mean BG of 126 versus 121 for those patients with high SB2M (≥6 μg/mL) (P < .08). Mean BG was 117 for the low albumin (≥3 g/dL) group and 128 for the high albumin group (<3 g/dL) (P < .03). Mean BG varied from 114 to 130 to 123 for patients grouped according to hemoglobin (<9.0 v 9.0 to 11.9 v ≥12.0 g/dL) (P < .01). Similar tests for differences based on clinical stage, monoclonal Ig, light chain, calcium, age, creatinine, and bone lesions were not significant.

Response and Enzyme Scores

As a single variable, there was a suggestion that pretreatment AP (<130 v ≥130) might be a significant predictor of complete and partial response (P = .05). However, in the multivariate analysis with all candidate prognostic variables, neither pretreatment AP nor pretreatment BG scores were significantly related to complete response.

Enzyme Scores and Survival

AP. The median survival did not differ significantly between those patients who had pretreatment AP scores available for analysis and those who did not. Initially, patients were grouped with AP scores such as less than 140, 140 to 179, 180 to 219, 220 to 259, and 260 or greater. These and other categorizations did not show a monotonic trend with survival. However, it was apparent that the patients in the lowest AP group had poorer survival than the other groups. Therefore, the group of patients having low AP values was examined further. When the patient population was divided into three groups with scores of less than 130, 130 to 229, and 230 or greater, median survival was significantly shorter for patients with low pretreatment AP values. The best separation of the population was at scores of less than 130 versus ≥130 (P = .001) (Fig 2).

BG. The median survival did not differ significantly between patients who had BG scores available for analysis and those who did not. Preliminary analysis of the BG scores suggested survival differences based on low, intermediate, or high scores. Patients with low scores (<100) appeared to have good survival, patients with intermediate scores (100 to 139) had the poorest survival, but the median
survival gradually improved as the scores increased from intermediate to high levels (≥140). Splitting patients with BG scores of less than 140 versus those with scores of 140 or greater suggested a difference in survival (P = .05) (Fig 3).

**Multivariate Analysis**

Patients were grouped based on AP and BG scores and they were analyzed in relation to SB2M, albumin, calcium, age, creatinine, hemoglobin, performance status, and myeloma stage.

**AP.** The subset of patients (15%) with a low AP score had significantly poorer survival, even when serum albumin, calcium, creatinine, SB2M, age, and hemoglobin were considered (P < .0001). Patients with low albumin or high calcium values had poor survival irrespective of AP scores. However, those patients with normal albumin or calcium values and high AP scores had better survival than those with low AP scores (Fig 4A and B). Patients with low creatinine values and high AP values had the best median survival, and those with high serum creatinine values and low plasma cell AP values had the poorest median survival (Fig 4C). The other two groups of patients had a similar median survival. With respect to SB2M and AP, the patients appear to fall into three groups: a favorable survival group (AP ≥130 and SB2M < 6), an intermediate group with either unfavorable AP of less than 130 or unfavorable SB2M of ≥6, and the poorest surviving group, in which AP and SB2M are both unfavorable (Fig 4D). Among both older and younger patients, low AP scores were associated with decreased median survival. Among
Table 2. Pretreatment Prognostic Variables and Survival in Myeloma

<table>
<thead>
<tr>
<th>Univariate analysis</th>
<th>P Value to Enter</th>
<th>Final Model P Value</th>
<th>Regression Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td>.29</td>
<td>.29</td>
<td>—</td>
</tr>
<tr>
<td>Low (&lt; 130) v high (≥ 130)</td>
<td>.001</td>
<td>.001</td>
<td>-.51</td>
</tr>
<tr>
<td>BG score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td>.18</td>
<td>.18</td>
<td>—</td>
</tr>
<tr>
<td>Low v intermediate v high</td>
<td>.12</td>
<td>.12</td>
<td>—</td>
</tr>
<tr>
<td>Low (&lt; 140) v high (≥ 140)</td>
<td>.05</td>
<td>.05</td>
<td>-.29</td>
</tr>
<tr>
<td>Multivariate analysis (variables in order of importance)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SB2M (log 10)</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>.90</td>
</tr>
<tr>
<td>Performance status (0-1 v 2-4)</td>
<td>.002</td>
<td>.001</td>
<td>.44</td>
</tr>
<tr>
<td>AP (≤ 130) v &gt; 130)</td>
<td>.002</td>
<td>.006</td>
<td>-.46</td>
</tr>
<tr>
<td>Age</td>
<td>.01</td>
<td>.01</td>
<td>.02</td>
</tr>
<tr>
<td>Stage (I-II v III)</td>
<td>.02</td>
<td>.02</td>
<td>.36</td>
</tr>
</tbody>
</table>

Patients with low hemoglobin values, median survival did not change significantly with varying AP scores. However, among patients with moderate and high hemoglobin, the patients with low AP scores had poorer survival.

BG. Patients with low SB2M values did not show any difference in median survival based on BG scores. However, patients with high SB2M and high BG scores had better survival than those with low BG scores (median 3.1 years v 1.6, respectively; P = .02).

Both enzymes. The candidate prognostic variables (both continuous values and indicators) of SB2M (log 10), age, albumin, calcium, creatinine, and hemoglobin, plus variables for AP, BG, sex, race, stage, and performance status were added into a log regression model. In the multivariate model, AP retained prognostic value whereas BG did not. The final set of prognostic variables chosen were SB2M (log 10), performance status, AP (< 130 v ≥ 130), age, and stage (Table 2).

DISCUSSION

Many observers have alluded to the fact that AP and BG levels are markedly increased in myeloma cells as compared with reactive plasma cells or plasma cells from patients with MGUS. Based on these observations there has been much interest in the use of AP or BG content of the plasma cells for differentiating nonmalignant versus malignant plasmacytic proliferations and as tumor markers in biologic studies of MM.

With the currently available cytochemical reagents the enzyme granularity for AP and BG was very strong and crisp. When we used the scoring method of Cassuto et al, almost all myeloma cases had scores in the range of 3.0 to 3.5 per cell, resulting in inadequate separation of cases based on enzyme quantity. Similar observations were made by Hoffman et al and Tartarollo et al. The method used in this study resulted in a broader range of scores so that different myeloma cases could be grouped and compared.

We were able to confirm that compared with plasmacytosis and MGUS there is a marked increase of AP and BG in MM. As observed by other investigators, we also found a significant overlap of enzyme scores between plasma cytosis, MGUS, and MM. When the enzyme scores are significantly high in a particular patient there is a higher probability that the plasmacytic proliferation is malignant. However, low or intermediate scores do not predict benign disease.

Because AP and BG are lysosomal acid hydrolases, it has been postulated that these enzymes could be involved in bone resorption and as a result they could contribute to the pathogenesis of osteolytic lesions in MM. Our study showed that AP scores were modestly higher in patients with bone lesions compared with those without bone lesions. However, neither AP nor BG were predictive of serum calcium levels or the extent of progression of bone lesions.

The most important finding was the discovery that the subset of patients with low AP have a uniformly poor prognosis (P = .001). In the multivariate analysis for survival prediction, an AP score of less than 130 was an important predictive factor (P = .006) following SB2M and performance status. As shown in Figs 2 and 4A through D, a low AP score was a strong predictor of poor survival and enhanced prognostic ability when applied in conjunction with other prognostic indicators such as albumin, calcium, creatinine, and SB2M. This counterintuitive finding raises several interesting questions about the pathophysiologic role of lysosomal enzymes in myeloma cells. Increased AP or BG enzyme activity are seen in activated histiocytes or T cells and T-cell malignancies. In contrast, AP and BG enzyme content is normal or decreased in B-cell malignancies such as chronic lymphocytic leukemia, Waldenström's macroglobulinemia, and non-Hodgkin's lymphoma. Thus, an increased quantity of AP or BG in myeloma is an anomaly in comparison to other B-cell neoplasms. Because lysosomal acid hydrolases are involved in the pathway of nucleoside, triphosphate, and purine metabolism, the increase in AP may be related to the high energy requirements for protein synthesis in myeloma. Alternatively, the increase might also be related to significantly increased cell death and autophagocytosis.

Various investigators have reported low AP activity in 10% to 22% of patients. Bataille et al theorized that low AP levels may indicate remission or less aggressive behavior. Hoffman et al followed-up on five myeloma patients longitudinally and they could not confirm Bataille's hypothesis. Our findings indicate that the low AP myeloma has an aggressive behavior with poor median survival. It is important to observe that the traditional myeloma prognostic factors, such as stage, SB2M, and performance status, convey information about the extent of tumor and/or the overall physical condition of the patient. Lysosomal enzyme scoring, on the other hand, deals with a direct biologic characteristic of the myeloma cells. Prospective studies of the low AP myeloma patients would be of interest to determine whether specific modulators could enhance AP expression or whether alternate therapeutic strategies might prove to be of greater value for these patients.
REFERENCES

Prognostic correlation of plasma cell acid phosphatase and beta-glucuronidase in multiple myeloma: a Southwest Oncology Group study [see comments]

SM Saeed, D Stock-Novack, R Pohlod, J Crowley and SE Salmon