1 STATISTICAL ANALYSIS

1.1 STUDY DESIGN

The study has a phase I/IIa design. The first phase I portion will be a standard 3+3 dose escalation study to include between 9 and 30 patients. However, it is highly unlikely that 30 patients will be enrolled (this would occur if dose expansion to 6 patients per level occurred at all of dose levels 0 through 4) and expected enrollment is expected to be 18 or fewer. The phase IIa portion of the study will expand the MTD cohort (or cohort 4 if MTD not reached) to enroll a total of 28 patients using a single stage design. Due to the 100 day lag between treatment and evaluation, an interim analysis for the primary endpoint is not feasible.

1.2 STUDY ENDPOINTS

1.2.1 PRIMARY ENDPOINTS

Phase I – The primary endpoint in phase I is the occurrence of a dose-limiting toxicity as defined in section 3.1.1.1.

Phase IIa – Rate of very good partial response (VGPR) + complete response (CR) in patients with relapsed MM treated with the melphalan + MTD of carfilzomib as determined in the phase I component of the study.

1.2.2 SECONDARY ENDPOINTS

Phase I and IIa – Pharmacodynamic effects of this combination at the proposed doses and schedule: changes in expression of Fanconi anemia/BRCA DNA repair genes and DNA fragmentation.

Phase IIa – Rate of overall response, defined as CR+VGPR+PR (Appendix B). Rate of PFS 12 months after AHSCT in the setting of carfilzomib + melphalan conditioning and carfilzomib maintenance therapy. Pharmacodynamic effect of maintenance regimens A and B assessed by proteasome inhibition in peripheral blood mononuclear cells. Subject's preference for regimens A and B. Measurement of proteasome inhibition is described in section 3.6.

1.2.3 SAFETY ENDPOINTS

Both phase I and phase IIa, transplant component – Engraftment kinetics, described by the median time for neutrophil and platelet engraftment (as defined in section 3.1.1.1) and the rate of engraftment by day 30. Frequency and nature of grades 3 and 4 non-hematologic AEs and SAEs.

Phase IIa, maintenance component – Frequency and nature of AEs and grades 3 and 4 AEs.

1.3 SAMPLE SIZE CONSIDERATIONS

The phase I sample size is determined based on the occurrence of DLTs at each dose level and the expansions. The minimum number of patients treated would be 6 (with 3 patients at each of dose levels 0 and -1) and the maximum would be 30 (with an expansion to 6 patients at each of dose levels 0 to 4). These are both unlikely and it is expected that the number of patients treated in the phase I portion will be between 12 and 18.

Patients will be enrolled in the phase 2 portion of the study to expand the MTD cohort (or cohort 4 if MTD is not reached) in order to reach 28 patients. There is little preliminary data to suggest a historical control response rate to the proposed treatment regimen. As a result, we are choosing to enroll 28 patients in order to be able to estimate the response rate with sufficient precision. More specifically, N=28 provides a half width of <0.20 for the 95% confidence interval for any response rate. This precision will provide sufficient information to determine if the regimen is sufficiently promising, in conjunction with the safety information learned in the phase IIa portion.

The phase IIa maintenance portion of the study is included to describe additional endpoints including proteasome inhibition in two different regiments of treatment, and patient preference. There will be no hypothesis testing performed and so there is no power calculation. No additional patients are enrolled specifically for the objectives addressed by the maintenance phase.

1.4 INTERIM AND SAFETY ANALYSIS

There is no interim analysis in Phase IIa for futility or efficacy. However, there will be continuous monitoring for safety. A sequential probability ratio test (SPRT) approach will be used. If there is strong evidence that the rate of grade 4 toxicities is 0.30 or above, as compared to a null rate of 0.15, the study will be stopped. The stopping boundary is based on the likelihood ratio comparing the null rate of 0.30 versus the alternative rate of 0.15 using binomial likelihoods. If the ratio favoring a rate of 0.30 (vs. 0.15) ever exceeds 10, then the study will be stopped. A likelihood ratio of 10 is similar to a significance level of 0.05. The stopping criteria for this approach are listed in Table 10 below.

Table 10. Stopping criteria for early stopping due to excessive toxicity. For example, if four of the first six patients experience a grade 4 toxicity, the study will be stopped. The last boundary (9 grade 4 toxicities in 28 patients) would not stop the study (because the maximum N is 28); but at the end of the study, the treatment would be deemed to toxic to take to the next phase of research.

Number of Grade 4 toxicities	Number of Patients Treated	Observed toxicity rate	Likelihood Ratio (favoring 30% toxicity rate)
4	6	67%	10.9
5	10	50%	12.1
6	15	40%	11.2
7	20	35%	10.3
8	24	33%	11.5
9	28	32%	12.8

1.5 PLANNED METHODS OF ANALYSIS

The primary objective of the phase I portion of the study is determination of the maximum-tolerated dose (MTD). This will be determined based on the algorithmic dose finding approach where the highest dose at which 0 or 1 of 6 patients experiences a DLT is designated as the MTD. The properties of this design are shown in Table 11 below where dose is escalated if either (a) 0 out of 3 evaluable patients experience a DLT, or (b) 1 out of 6 patients experience a DLT.

Table 11: Probability of escalation based on true DLT rates ranging from 10% to 60%. We assume that DLT rates of 30% and below are acceptable. For DLT rates that are 40% or greater, we have a relatively low chance of escalation, especially when the DLT rate is as high as 50% or 60%.

True DLT rate	Probability of escalation
10%	0.91
20%	0.71
30%	0.49
40%	0.31
50%	0.17
60%	0.08

The primary objective of the phase IIa portion of the study is to estimate the response rate. This will be done by estimating the proportion of patients who experience a response (defined above) with its 95% confidence interval.

The secondary objectives of the study include describing toxicities in both the phase I and IIa portions, and describing the pharmacodynamics effects of the treatment in phase I and phase IIa. This will be done by tabulating toxicities by type and grade in each phase. Pharmacodynamic measures will be quantified by changes in quantitative measures (e.g. expression) between baseline and follow-up. These will be summarized using graphical displays and summary statistics. PFS will be graphically displayed using Kaplan-Meier curves and 12 month PFS will be estimated with its 95% confidence interval.

For those patients in the maintenance phase of the phase IIa portion of the study, proteasome inhibition will be treated as a continuous variable and summarized by treatment arm using graphical displays and summary statistics. Preference for A vs. B and completion rates will be estimated by proportions with exact 95% confidence intervals. Given that the number of patients in the maintenance portion of the study is expected to be small (20-22), hypothesis

testing will not be performed to compare patients who receive A first versus those who receive B first.