The UACC812 data was split across two RPPA experiments. Each batch contains different inhibitors, but both batches contain DMSO.

Below is a description of how the two UACC812 batches were combined and normalized to obtain a single data set for UACC812. The starting point is the 'Normalized Linear Value' data (on log2 scale) for each batch[1].

1) Any antibodies not included in both batches are removed

2) The following is carried out for each antibody:
Using log2 'Normalized Linear Value' data, the mean and standard deviation of the DMSO samples in each batch are calculated, giving values μ_1,σ_1 and μ_2,σ_2 for batch 1 and batch 2 respectively.
Note that, for each batch, there are 16 replicates for DMSO,0min (all other DMSO conditions consist of a single replicate). These 16 replicates are averaged prior to calculating the mean and standard deviation across the DMSO samples in a batch.

3) All the samples in batch 2 are then re-scaled/re-centred so that the mean and standard deviation of the DMSO samples in batch 2 (after replicate averaging) agree with the batch 1 quantities (μ_1 and σ_1).
i.e. a sample in batch 2 with value x becomes

$$\mu_1 + \frac{\sigma_1(x - \mu_2)}{\sigma_2}.$$

This re-scaling/re-centring is applied to each individual replicate and not to replicate-averaged data.

4) The two batches are then combined to get a single data set for UACC812. The data is provided on a linear scale (not log2 scale).

This batch-normalized data is provided in CSV files *UACC812_main.csv* and *UACC812_full.csv,* and also in MIDAS formatted files M*D_UACC812_main.csv* and *MD_UACC812_full.csv*.
Note that all individual replicates are provided in these data (we have not averaged any replicates, either within a batch or across the batches; replicate-averaging in step 2 is performed solely for the purpose of calculating μ_1,σ_1,μ_2, and σ_2).

The original 'Normalized Linear Value' data for the two individual batches are also provided in the folder *extra_UACC812_files* in the CSV format.

[1] The 'Normalized Linear Value' data is obtained from raw 'SuperCurve' data by applying the protein loading normalization procedure developed by the RPPA Core Facility at MD Anderson Cancer Center. For details see question 4 on this FAQ page:
http://www.mdanderson.org/education-and-research/resources-for-professionals/scientific-resources/core-facilities-and-services/functional-proteomics-rppa-core/faq/functional-proteomics-reverse-phase-protein-array-core-facility-faq.html