I have now run the quantification for all of the 226 samples I have.
You can just work with the first 1 or 2 files for now, we can add the rest later on.

The results of these are now in .csv files (the values are separated by commas, rather than tabs)
There is a value in the 'intensity' column, which we can use for scoring.
Many of these are blank - we can restrict to only using the peptides that have a value in them.

In each individual file, see how many peptides are in each protein.

? from .fasta file in the first version of data

If there is just 1 peptide, ignore it (But keep a count of how many are being ignore, we may want to add these back in)

Make a 'protein total' which is the just the sum of the 3 peptides with the largest intensity values

? sum of intensity value ? for each protein

(Or 2 if there are only 2)

Score the peptide as 100 if it has an intensity value of more than 20% of the protein total

Score the peptide as 50 if it has an intensity value of less than 20% of the protein total

And you already know to score the peptide as 0 if it is not present, but the parent protein is.

For each file, you will then have a score for each peptide.
We would then need to merge all of the output files, and come up with a final list of all peptides, with a score for each.
To merge the score, just use the most common value found.

I can work up some examples, but that would be a few days, as I have other things I need to do this week.

To summarise:
Use these .csv files in place of the .csv files you used previously
For each file, score each peptide as 100,50,0 and write out the results
Merge the results to get a single file with each peptide and a score
Afterwards:
Convert peptides to 'chemical information' (I will be looking into this)
Run chemical information through Random Forests to look for interesting results