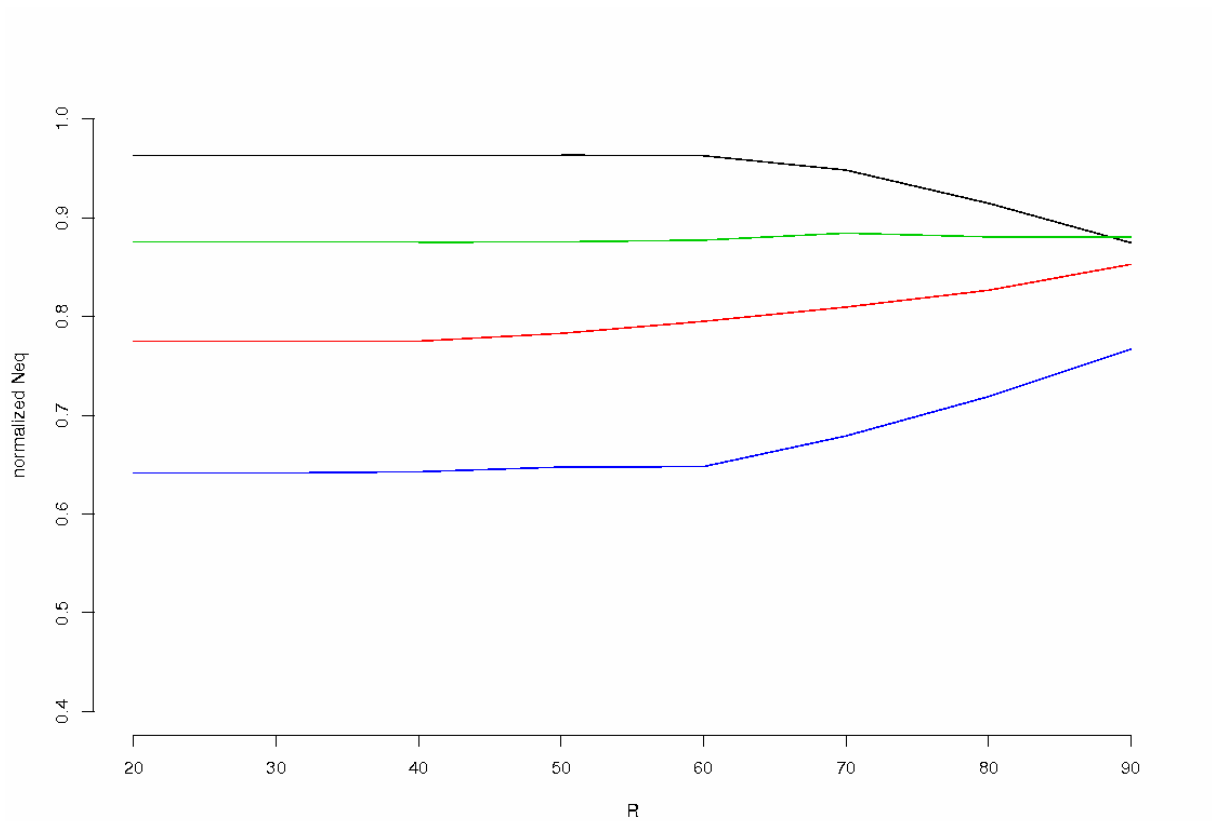


Supplementary material 1:



The four different evolution of PU sizes during cutting. The N_{eq} profile of each protein of the non-redundant databank has been computed (R ranging from 20 to 90, see Figure 3). The profiles have been clustered into four groups. Three are close while the fourth has a particular behavior (see text).