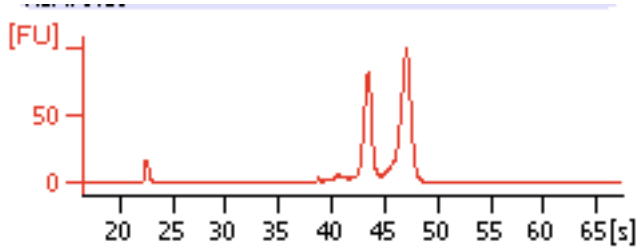
Here are Bioanalyzer traces from various points in my illumina library preparation protocol for RNA-seq, modified to sequence only 3' transcript ends. Graphs are from different samples since limits in starting amounts don't allow me to validate any singe sample at each of its steps.

My question is: **what are the high molecular weight molecules between 1000-4000bp in my cDNA samples (graphs D1 & D2)?** I would think that if the mRNA template is fragmented, the lack of 1-4k length rna templates (avg fragment length <200bp, graph C) would preclude synthesis of long fragments. And RNase H treatment in my protocol should eliminate long RNA contamination. It looks like there are some cDNAs of appropriate length (i.e. <200bp). Thanks for any ideas!

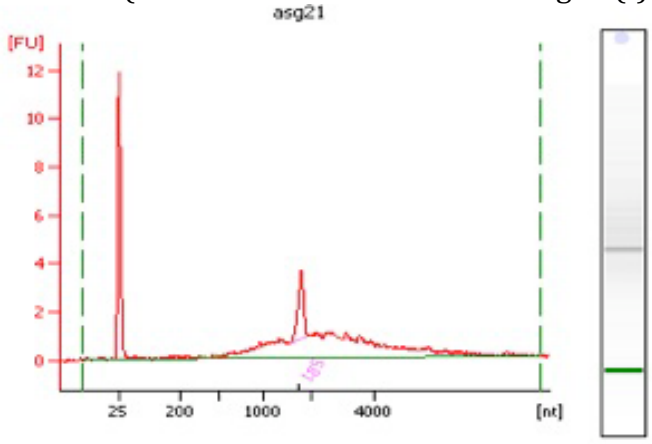
The simplified protocol is:

1. From total RNA (graph A, below), isolate mRNA (graph B) using oligo d(t) dynabeads
2. Fragment mRNA using Ambion RNA fragmentation reagents (graph C)
3. Select poly-A fragments from fragmented mRNA using oligo d(t) dynabeads
4. Synthesize first strand cDNA using oligo d(t) primers and Invitrogen Superscript II
5. Synthesize second strand cDNA using DNA pol I and RNase H (graphs D1 & D2)
6. Blunt-ending, adenylation, adapter-ligation, size-selection, pcr enrichment, etc

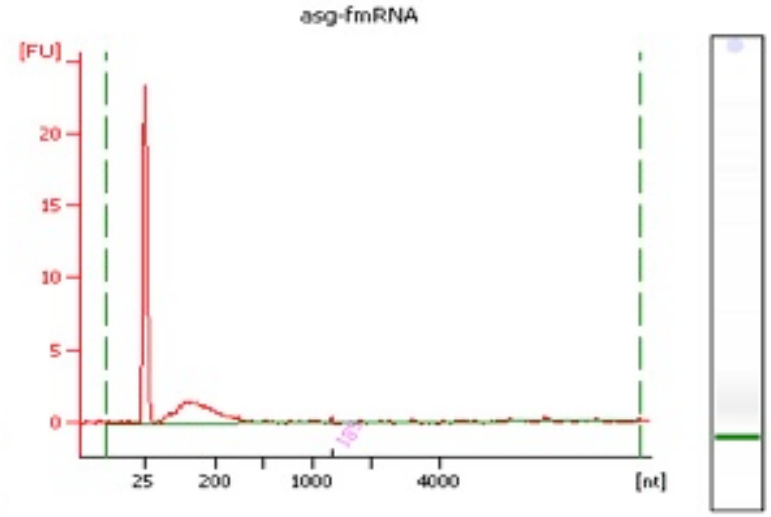
A. Total RNA:



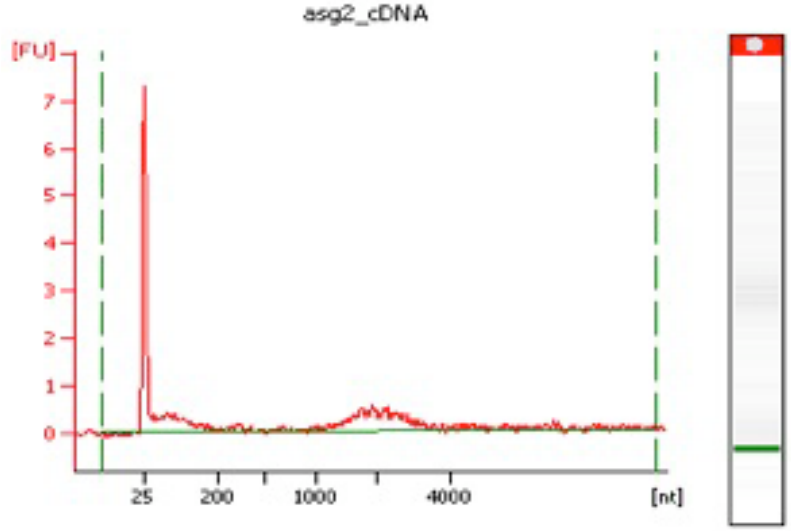
B. mRNA (after double treatment with oligo d(t) dynabeads according to standard protocol):



C. mRNA fragmented for 5m at 70 C using Ambion fragmentation reagents:



D1. Double-stranded cDNA synthesized from fragmented mRNA



D2. double-stranded cDNA synthesized from oligo d(t)-selected fragmented RNA: