Supplemental Outline of Tables

1. Table S5 Back-to-back CID/HCD Experiment (Fig. 1)
   Tab 1 Column headings description
   Tab 2 Total phosphopeptides identified by CID
   Tab 3 Unique sites identified by CID
   Tab 4 Total phosphopeptides identified by HCD
   Tab 5 Unique sites identified by HCD
   Tab 6 Total phosphopeptides identified by both CID and HCD
   Tab 7 Unique sites identified by both CID and HCD
   Tab 8 Proteins identified by CID
   Tab 9 Proteins identified by HCD
   Tab 10 Proteins identified by both CID and HCD

2. Table S6 pY enrichment experiment (Fig. 2)
   Tab 1 Column headings description
   Tab 2 Total phosphopeptides identified by CID
   Tab 3 Unique sites identified by CID
   Tab 4 Total phosphopeptides identified by HCD
   Tab 5 Unique sites identified by HCD
   Tab 6 Total phosphopeptides identified by both CID and HCD
   Tab 7 Unique sites identified by both CID and HCD
   Tab 8 Proteins identified by CID
   Tab 9 Proteins identified by HCD
   Tab 10 Proteins identified by both CID and HCD

3. Table S7 SCX/IMAC experiment (Fig. 3)
   Tab 1 Column headings description
   Tab 2 Total phosphopeptides identified by CID
   Tab 3 Unique sites identified by CID
   Tab 4 Total phosphopeptides identified by HCD
   Tab 5 Unique sites identified by HCD
   Tab 6 Total phosphopeptides identified by both CID and HCD
   Tab 7 Unique sites identified by both CID and HCD
   Tab 8 Proteins identified by CID
   Tab 9 Proteins identified by HCD
   Tab 10 Proteins identified by both CID and HCD

4. Table S8 Back-to-back high resolution CID/HCD Experiment (Fig. S3)
   Tab 1 Column headings description
   Tab 2 Total phosphopeptides identified by high resolution CID
   Tab 3 Unique sites identified by high resolution CID
   Tab 4 Total phosphopeptides identified by HCD
   Tab 5 Unique sites identified by HCD
   Tab 6 Total phosphopeptides identified by both high resolution CID and HCD
   Tab 7 Unique sites identified by both high resolution CID and HCD
   Tab 8 Proteins identified by high resolution CID
   Tab 9 Proteins identified by HCD
   Tab 10 Proteins identified by both CID and HCD
5. Table S9 SCX/IMAC experiment (Table S1)
   Tab 1 Column headings description
   Tab 2 Total phosphopeptides identified by CID
   Tab 3 Unique sites identified by CID
   Tab 4 Proteins identified by high resolution CID

6. Table S10 Back-to-back CID/HCD Experiment (Table S2)
   Tab 1 Column heading description
   Tab 2 Total phosphopeptides identified by CID (MASCOT)
   Tab 3 Unique sites identified by CID (MASCOT)
   Tab 4 Total phosphopeptides identified by HCD (MASCOT)
   Tab 5 Unique sites identified by HCD (MASCOT)
   Tab 6 Total phosphopeptides identified by both CID and HCD (MASCOT)
   Tab 7 Unique sites identified by both CID and HCD (MASCOT)
   Tab 8 Proteins identified by CID (MASCOT)
   Tab 9 Proteins identified by HCD (MASCOT)
   Tab 10 Proteins identified by both CID and HCD (MASCOT)

7. Table S11 Reanalysis of data from Mann and co-workers(23) (Table S3)
   Tab 1 Column headings description
   Tab 2 Total phosphopeptides identified by CID1 (MASCOT)
   Tab 3 Unique sites identified by CID1 (MASCOT)
   Tab 4 Total phosphopeptides identified by HCD1 (MASCOT)
   Tab 5 Unique sites identified by HCD1 (MASCOT)
   Tab 6 Total phosphopeptides identified by HCD2 (MASCOT)
   Tab 7 Unique sites identified by HCD2 (MASCOT)
Table S1. Comparison of the data sets from 10 SCX fractions analyzed using either the OrbiVelos platform with HCD- or CID-type fragmentation or Orbi XL platform with CID-type fragmentation. Matches to reversed (decoy) sequences are in parentheses. CID-based spectra were always detected in the ion trap whereas HCD spectra were detected in the orbitrap.

<table>
<thead>
<tr>
<th>Type (Runs)</th>
<th>Total MS/MS</th>
<th>Total Peptides</th>
<th>Unique sites</th>
<th>Peptide FDR</th>
<th>Proteins</th>
<th>Protein FDR</th>
<th>Analysis time</th>
</tr>
</thead>
<tbody>
<tr>
<td>OrbiVelos CID (10)</td>
<td>266,304</td>
<td>53,679</td>
<td>17,909</td>
<td>0.17% (90)</td>
<td>4,871</td>
<td>1.01% (49)</td>
<td>13.3</td>
</tr>
<tr>
<td>OrbiVelos HCD (10)</td>
<td>115,427</td>
<td>29,326</td>
<td>11,925</td>
<td>0.26% (76)</td>
<td>3,929</td>
<td>0.99% (39)</td>
<td>13.3</td>
</tr>
<tr>
<td>Orbi XL CID (10)</td>
<td>197,033</td>
<td>38,871</td>
<td>11,928</td>
<td>0.12% (48)</td>
<td>3,483</td>
<td>0.95% (34)</td>
<td>13.3</td>
</tr>
</tbody>
</table>

Table S2. Comparison of Sequest and MASCOT search algorithms using the data shown in Figure 1. Back-to-back HCD and CID analysis on the same precursor ions were collected. Very similar search and post-processing parameters were used. Both search algorithms performed similarly for HCD- and CID-type MS/MS spectra. The fragment ion tolerance for the HCD searches was 0.02 Da for both Sequest and MASCOT. False discovery rates were to set to 1% at the protein level for each search. Matches to reversed (decoy) sequences are in parentheses.

<table>
<thead>
<tr>
<th>Type</th>
<th>Total MS/MS</th>
<th>Total Peptides</th>
<th>Unique sites</th>
<th>Peptide FDR</th>
<th>Proteins</th>
<th>Protein FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>CID Sequest</td>
<td>16,640</td>
<td>6,284</td>
<td>3,326</td>
<td>0.13% (8)</td>
<td>1,782</td>
<td>0.95% (7)</td>
</tr>
<tr>
<td>CID MASCOT</td>
<td>16,640</td>
<td>5,715</td>
<td>3,027</td>
<td>0.28% (16)</td>
<td>1,446</td>
<td>0.97% (14)</td>
</tr>
<tr>
<td>HCD Sequest</td>
<td>16,640</td>
<td>7,692</td>
<td>3,878</td>
<td>0.16% (12)</td>
<td>1,591</td>
<td>0.92% (8)</td>
</tr>
<tr>
<td>HCD MASCOT</td>
<td>16,640</td>
<td>7,376</td>
<td>3,555</td>
<td>0.25% (18)</td>
<td>1,616</td>
<td>0.99% (16)</td>
</tr>
</tbody>
</table>

Table S3. Phosphopeptide analysis of .RAW files from reference 25. All .RAW files were downloaded and searched by MASCOT with the same parameters as found in Table S2. Similar post-processing parameters were used as in all our experiments where the FDR was set to 1% at the protein level. The Ascore algorithm was used to calculate the unique sites found in all three experiments.
Supplementary Table 4. Phosphopeptide analysis by a back-to-back, alternating high resolution CID- and HCD-type fragmentation for same-precursor ions. The 5 most abundant ions from each full MS cycle were subjected to sequential CID (orbitrap detection) and HCD (orbitrap detection) fragmentation. The table summarizes peptide and protein identifications from this experiment. Matches to reversed (decoy) sequences are shown in parentheses.
Supplementary Figure 1. Typical example spectra for CID- and HCD-type analyses in Figure 1. (A) CID spectrum matched to the peptide, EGpSPAPLEPEPGASQPK from protein Foxk2. A prominent neutral loss of phosphoric acid is seen, but b- and y-type fragment ions still facilitated the database match. (B) HCD spectrum from the same precursor mass as the CID spectrum. No loss of phosphoric acid is seen, and the low mass region is present. Xcorr, ΔCn’, and Ascore values were all higher for the HCD spectrum for this peptide. Note that all spectra can be visualized via hyperlinks in the supplementary tables.
Supplementary Figure 2. Examples of CID and HCD fragment spectra for a doubly-charged, phosphotyrosine-containing peptide from the experiment performed in Figure 3. (A) CID spectrum collected in the ion trap and matched to the sequence LENLVYpYNR from Camkv. (B) HCD spectrum collected in the orbitrap and matched to the same peptide. The low mass range is intact and contains the phosphotyrosine-specific immonium ion at 216.0415. All MS/MS spectra are available through hyperlinks in the supplementary tables. ΔCn' and Ascore values were higher for the HCD spectrum for this peptide. All spectra can be visualized via hyperlinks in the supplementary tables.
Supplementary Figure 3. Global data-dependent settings for CID (A) and HCD (B) analyses in Nagaraj et al. JPR 2010 (25). Xcalibur .RAW files acquired by Nagaraj and co-workers were downloaded as supplementary data from ProteomeCommons.org at the following hash: /Gyf6Csx8Xi8xaUTof4/OcFDvL3Td6j4UPLceZSTXL2kdZr90Ub5j6NdIduK6+ehqHJ37d9GZSQTKaDtUM4/gMsNYAAAAAAAATLQ==. Each .RAW file was examined for global data-dependent settings as shown. The instrument settings for all CID .RAW files were the same as shown in Panel A, which is a screen capture of their methods embedded in the .RAW file shown. All HCD .RAW files were acquired using the settings in Panel B. For CID analyses, the exclusion mass window was set to exclude a 15 m/z window around the selected precursor ion. However, for HCD analyses, this window was set at ±10 ppm. For a peptide of m/z 1,000 this is a 750-fold difference in the size of this window, which was in effect for up to 60 seconds. During a shotgun sequencing experiment, this large window could effectively block out most of the mass spectrum from further analysis. This difference in parameters likely influenced their finding that HCD fragmentation of phosphopeptides produced more identifications than CID.

Finally, there were other potentially important differences between the reported and actual instruments settings as follows (setting, HCD, CID): Threshold for triggering MS/MS, 10000, 1000; Dynamic exclusion duration, 120 s, 60 s; Precursor isolation width, 4 m/z, 2 m/z.
Supplementary Figure 4. Evidence of systematic over-counting of phosphorylation sites in Nagaraj et al. (2010) (25). To better understand how phosphorylation sites were inferred from identified phosphopeptides, we examined phosphorylation site lists that were provided as supplementary materials accompanying the manuscript: pr100637q_si_002.txt, pr100637q_si_003.txt, and pr100637q_si_004.txt. Each list summarizes the phosphorylation sites identified in each experiment and indicates the single best matching MS/MS scan for each site. Importantly, close examination of the scans assigned to each site demonstrates systematic over-counting of phosphorylation sites. Two representative examples are provided in A. The first of these examples is scan 6269, which matches a singly-phosphorylated peptide. While the localization of this single phosphorylation may be uncertain, it should account for at most one phosphorylation event. However, this single spectrum is cited as evidence to support four different phosphorylation sites at positions 645, 646, 647, and 649 (indices indicate the ID assigned to each site in the supplementary Table). Similarly, although scan 21154 matches a doubly phosphorylated peptide, it is simultaneously used to justify seven distinct phosphorylation sites. These are not isolated examples. Part B summarizes for all 12,000 scans listed in supplementary Table pr100637q_si_002 the number of phosphates borne by peptides matching each spectrum (vertical), versus the number of distinct phosphorylation sites that have been assigned to each spectrum (horizontal). Each number in the array indicates the number of MS/MS scans within each bin. Bins highlighted in grey indicate scans for which the number of assigned phosphorylation sites matches the number of phosphates borne by the matching peptide, while bins highlighted in red indicate scans for which the number of assigned phosphorylation sites is less than the number of phosphates borne by the matching peptide; both of these scenarios could...
be correct. However, highlighted in green are scans that were used to justify numbers of phosphorylation sites that exceed the total number of phosphates borne by their matching peptides. For example, there are 574 scans assigned to singly phosphorylated peptides that were used to support the existence of two distinct phosphorylation sites each. In one extreme case, a singly phosphorylated peptide was simultaneously cited as evidence for eleven different phosphorylation sites. Part C indicates the extent to which this phenomenon was observed in each of three supplementary tables. Overall, between 4 and 11% of spectra were used to account for more sites of modification than their sequences contained. Numbers of excess sites listed in each supplementary table ranged from 500 to over 2300 sites, accounting for up to 14% of reported phosphorylation events. It is important to note that this analysis provides only a minimum estimate of the extent to which over-counting inflates the numbers of reported sites. Identifying all such cases requires knowledge of all sites assigned to each MS/MS scan. Because complete lists of peptides and the sites they bore were not provided, we can only identify over-counting of sites in relatively rare cases where the same MS/MS scan provided the best observations of several scans.
Supplementary Figure 5. Phosphopeptide distributions by charge state for the data in Figures 1-3. The tables summarize peptide and protein identifications from each experiment for given charge states. For each experiment, phosphopeptide matches were set to a 1% FDR at the protein level.
Supplementary Figure 6. Multivariate Analysis of Phosphopeptide Identifications. To identify characteristics that distinguish phosphopeptides identified by HCD alone (Blue), CID alone (Red), or both strategies (Black), we profiled all phosphopeptides identified in back-to-back HCD-CID analyses with respect to 49 different properties. Because CID and HCD spectra were collected sequentially for each precursor during data-dependent acquisition, differences reflect only preferences in peptide identification. Each histogram displays proportions of peptide ID's observed for each set of phosphopeptides. Features include mass, charge, peptide length, and measures of hydrophobicity. The total number of amino acids in each peptide, as well as counts of each amino acid within each peptide sequence (e.g. 'Alanine Count') are also included. The fractional occurrence of each amino acid is included as well (e.g. 'Alanine Fraction': the number of alanines divided by the total amino acid count). Overall, there is little or no difference observed among peptides identified by either fragmentation approach with respect to any of the parameters considered. The final counts were as follows: CID-only (181); HCD-only (1604); CID + HCD (6122). The CID-only lines often appear jagged because they consist of smaller numbers (181 events). Finally, even though considered as HCD-only or CID-only, 57 and 55%, respectively, of these events were also identified by the alternate fragmentation method using Sequest, but they failed to pass subsequent thresholding to a protein FDR of 1%. 
Supplementary Figure 7. Investigation of an HCD-CID Decision Tree Method. If it were possible to predict prior to MS/MS acquisition which features would be most effectively analyzed via CID versus HCD, such information could be used to maximize peptide identifications in data-dependent experiments. Previously, decision tree methods have been used to optimize application of CID versus ETD (42) as well as ETD versus HCD (43) for proteomics applications. We partitioned by charge state those features selected for MS/MS analysis during independent CID and HCD analyses and measured the probability of successful phosphopeptide identification as a function of m/z (42) (top row). Features were assigned to 50-Thompson bins for probability determination. For comparison, numbers of features in each bin are displayed as well (bottom row). Though HCD generally displays higher probabilities of success throughout, each scan takes more than twice as long as an equivalent CID scan (HCD average: 0.303 sec; CID average: 0.136 sec). Once probabilities are adjusted to account for scan time, CID is preferred for 2+, 3+, and 4+ ions regardless of mass-to-charge ratio (middle row). Though HCD performs comparably, or perhaps slightly better for the small number of 5+ ions present in these samples, the difference is negligible. These probability distributions suggest that there would be no benefit to a decision tree method choosing between these two fragmentation strategies, as low resolution CID would always be the method of choice, regardless of charge state and mass-to-charge ratio.