

# Electron Transfer Dissociation and Multi-Stage Activation Analysis of Human Kinase Sites of Phosphorylation

Martin P. Hornshaw<sup>1</sup> and Nick Morrice<sup>2</sup>

<sup>1</sup>Thermo Fisher Scientific, Hemel Hempstead, UK and <sup>2</sup>MRC Protein Phosphorylation Unit, Dundee, UK

## Overview

**Purpose:** Assess the utility of different dissociation techniques for phosphorylation site analysis of human kinases.

**Methods:** Comparison of neutral loss based techniques with electron transfer dissociation during LC/MS.

**Results:** The different dissociation techniques were found to be entirely complementary and, rather surprisingly, a number of novel phosphorylation sites were identified on such well known kinases as PKB. ETD data on average identified more sites of phosphorylation and with greater discrimination relative to CID based approaches.

## Introduction

The analysis of phosphopeptides on a linear ion trap mass spectrometer has typically been performed by Data Dependent™ MS/MS followed by MS<sup>3</sup> of the putative neutral loss peak if observed in the MS/MS spectrum (neutral loss MS<sup>3</sup>). The strength of this approach is the sensitivity in full scan MS<sup>n</sup> of linear ion trap technology and the characteristically strong neutral loss peak serving to flag the MS/MS spectrum as potentially belonging to a phosphopeptide. However, two novel techniques have recently been developed that can be applied to the analysis of protein phosphorylation. Here we seek to better understand the utility of multi-stage activation (MSA)<sup>1</sup> and electron transfer dissociation (ETD)<sup>2</sup> applied to the analysis of sites of phosphorylation. Illustrations of the MSA and ETD processes are shown below in figure 1.

FIGURE 1a. Illustration of Multi-Stage Activation Scan

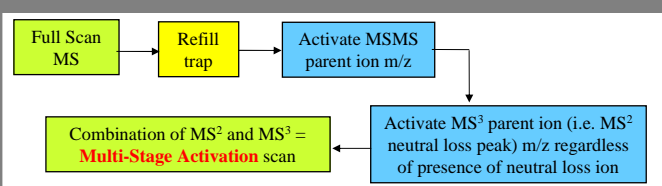
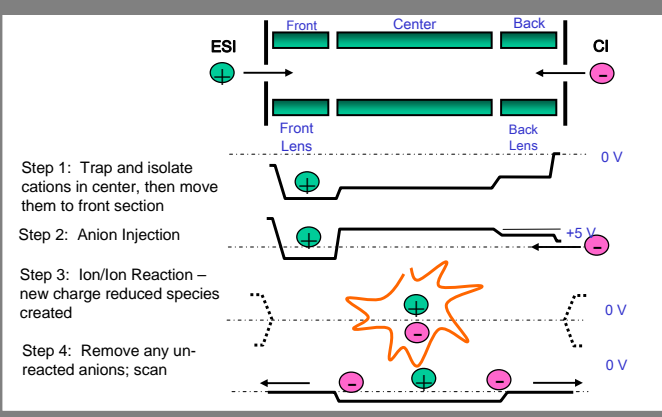


FIGURE 1b. Illustration of Electron Transfer Dissociation Process



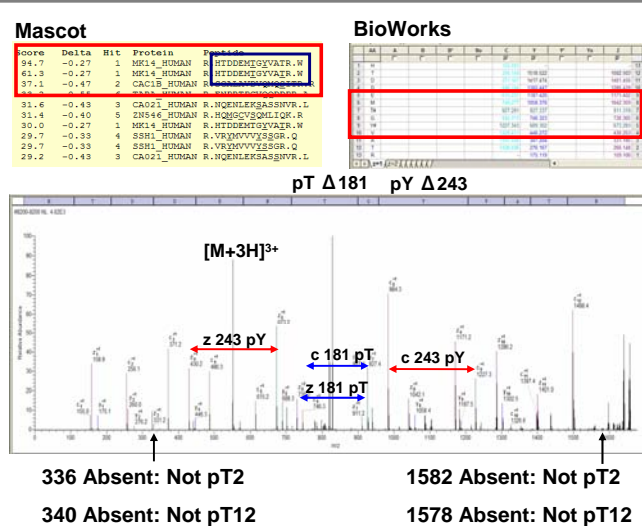
## Methods

Recombinant human kinases were studied. The proteins were reduced, alkylated and digested. The samples were analyzed by LC-MS<sup>n</sup> with nanoflow-reverse phase LC. The multi-stage activation and Data Dependent neutral loss MS<sup>3</sup> experiments were performed on both the Thermo Scientific LTQ Orbitrap™ and LTQ XL™ instruments. ETD experiments were performed exclusively on an LTQ XL. Only data from the LTQ XL is further discussed in this work. The typical workflow in the LTQ XL was to perform alternating CID (MSA) and ETD for the same precursor where precursors were selected in a Data Dependent manner following an enhanced resolution scan to enable charge state determination. In addition, 2+ precursor ions, when ETD was performed, were further activated by supplementary activation, that is, electron transfer dissociation was performed but unfragmented the precursor was then further activated by CID. Database searching was performed with SEQUEST® residing within BioWorks 3.3.1 software and with Mascot 2.1 and 2.2 (Matrix Science) against all human entries in Swiss-Prot or against a proprietary phosphoprotein database (MRC).

## Results

An example of the benefits of LC-MS using ETD for the identification of sites of phosphorylation is shown in figure 2. A doubly phosphorylated peptide from SAPK2a searched against the Swiss-Prot database produces a highly confident identification with both Mascot and SEQUEST. Note, no neutral loss was observed, which is typical for ETD. The difference between the best scoring ID and the second best using Mascot is 33.4. In addition, manual examination of the ETD MS/MS spectrum shows absence of evidence for phosphorylation at other potential sites of phosphorylation. The peptide is phosphorylated at a tyrosine and a threonine residue. Mass differences between ions of 243 and 181 clearly indicate the presence of phosphothreonine and phosphotyrosine at residues 7 and 9 respectively.

FIGURE 2. Example of Identification of Sites of Phosphorylation using ETD. Database searching was performed with Mascot and SEQUEST in BioWorks. Difference in MOWSE score for best and second best ID is significant. In addition, manual examination of the data revealed absence of evidence for phosphorylation at other potential sites.



For PKB, the ETD data was thoroughly analyzed to see which phosphopeptides were present and what the quality of identification was. A number of novel phosphopeptides were observed with confident identifications (figure 3 and table 1).

FIGURE 3. Sequence coverage of PKB analyzed by ETD. A number of novel phosphopeptides were observed (in blue).

1 MSYHHHHHHH DYDIPTTENL YFGQMGMSMD FRSGSPSDNS GAEEMEVS LA  
51 KPKHRVTMNE FEYLKLLGK TFGKVLVKE KATGRYAMK ILKKEVIVAK  
101 DEVAHTLTEN RVLQNSRHFP LTALKYSFQT HDRLCFVMEY ANGGELFFHL  
151 SRERVFSEDR ARFYGAEIVS ALDYHLSEKN VVYRDLKLEN LMLDKDGHK  
201 ITDFGLCKEG IKDGATMKTFCGTPPEYLAPEVLEDNDYGR VDWVWGLGVVM  
251 YEMMCGRLPF YNQDHEKLFELIMEEIRFP RTLGPEAKSL LSGLLKDPK  
301 QRLGGGSEDA KEIMQHRFFA GIWVQHYVEK KLSPPFPKQV TSETDTR YFD  
351 EEFQAQMITI TPPDQDSME CVDSERRPHF PQFDYSASGT A

ETD: 93% coverage (shown in red)

Black indicates that the peptide has been previously observed while blue indicates that this has not been previously observed.

SGSPSDNSGAEEMEVS LAKPK Observed as un-, mono-, di- and tri-phosphorylated  
HRVTMNEFEYLKLLGK HRVTMNEFEYLK HRVTMNEFEYLKLLGKGTGFK  
ILKKEVIVAKDEVAHTLTENR KEVIVAKDEVAHTLTENR EVIVAKDEVAHTLTENR  
ITDFGLCKEGIKDGATMKTFCGTPPEYLAPEVLEDNDYGR  
EGIKDGATMKTFCGTPPEYLAPEVLEDNDYGR Observed as mono- and diphosphorylated  
TLGPEAKSLSGLLK TLGPEAKSLSGLLKDPK  
KLSPPFPKQVTSETDTR

TABLE 1. Phosphopeptides Identified by ETD and MSA for PKB.

CID and ETD IDs are Mowse scores of best and second best peptide identifications. The difference between 1<sup>st</sup> and 2<sup>nd</sup> Scores and total number of phosphopeptides identified is used as an indication of the ability of ETD and MSA to identify sites of phosphorylation.

| Peptide Sequence                            | CID (MSA) 1 <sup>st</sup> and 2 <sup>nd</sup> IDs |                 | ETD 1 <sup>st</sup> and 2 <sup>nd</sup> IDs |                 |
|---|---|-----------------|---|-----------------|
|   | 1 <sup>st</sup>                                   | 2 <sup>nd</sup> | 1 <sup>st</sup>                             | 2 <sup>nd</sup> |
| SGSPSDNSGAEEMEVS LAKPK                      | 107   | 13              | 122   | 11              |
| SGSPSDNSGAEEMEVS LAKPK+p                    | 126   | 120             | 107   | 107             |
| SGSPSDNSGAEEMEVS LAKPK+2p                   | 101   | 94              | 108   | 108             |
| SGSPSDNSGAEEMEVS LAKPK+3p                   | 74  | 63              | 45  | 45              |
| HRVTMNEFEYLK                                | 65  | 22              | 85  | 13              |
| HRVTMNEFEYLK+p                              | 68  | 31              | 91  | 6               |
| HRVTMNEFEYLKLLGK                            | 74  | 25              | 119   | 25              |
| HRVTMNEFEYLKLLGK+p                          | 66  | 32              | 101   | 28              |
| HRVTMNEFEYLKLLGKGTGFK                       | 76  | 15              | 98  | 12              |
| HRVTMNEFEYLKLLGKGTGFK+p                     | ND  | ND              | 70  | 23              |
| EVIVAKDEVAHTLTENR                           | 121   | 13              | 98  | 22              |
| EVIVAKDEVAHTLTENR+p                         | 94  | 90              | 79  | 67              |
| KEVIVAKDEVAHTLTENR                          | 142   | 19              | 99  | 28              |
| KEVIVAKDEVAHTLTENR+p                        | 54  | 47              | 81  | 57              |
| ILKKEVIVAKDEVAHTLTENR+p                     | 70  | 67              | 81  | 62              |
| EGIKDGATMKTFCGTPPEYLAPEVLEDNDYGR+p          | 82  | 82              | 102   | 76              |
| EGIKDGATMKTFCGTPPEYLAPEVLEDNDYGR+2p         | 71  | 55              | 82  | 56              |
| ITDFGLCKEGIKDGATMKTFCGTPPEYLAPEVLEDNDYGR+p  | 76  | 75              | 67  | 41              |
| ITDFGLCKEGIKDGATMKTFCGTPPEYLAPEVLEDNDYGR+2p | 67  | 65              | 56  | 49              |
| TPCGTPPEYLAPEVLEDNDYGR+p                    | 87  | 85              | 27  | 27              |
| TLGPEAKSLSGLLK+p                            | 76  | 50              | 93  | 49              |
| TLGPEAKSLSGLLKDPK+p                         | 39  | 34              | 87  | 58              |
| KLSPPFPKQVTSETDTR+p                         | 24  | 17              | 54  | 28              |
| RPHFPQFDYSASGTA+p                           | 68  | 63              | 27  | 24              |

1<sup>st</sup>-2<sup>nd</sup>/1<sup>st</sup> and 1<sup>st</sup>-2<sup>nd</sup> CID 0.16 and 10.2 (n=17) ETD 0.305 and 24.8 (n=18)

A very simple attempt was made to quantify the difference between MSA and ETD (see tables 1 and 2 for examples) with respect to characterizing sites of phosphorylation. The best and second best Mowse scores after database searching with Mascot were used to indicate the quality of the identification of the phosphopeptide and the confidence of correct location of the site of phosphorylation. Typically, if more than one potential site of phosphorylation exists, then the best and second best scores after database searching will be for the same peptide sequence but with different potential sites of phosphorylation. ETD data, on average, was better able to discriminate between potential sites of phosphorylation than was MSA data. In addition, more phosphopeptides were detected with ETD than with MSA. Lastly, a number of novel phosphorylation sites were observed (see figure 3).

TABLE 2. Phosphopeptides Identified by ETD and MSA for MST2.

CID and ETD IDs are MOWSE scores of best and second best peptide identifications. The difference between 1<sup>st</sup> and 2<sup>nd</sup> Scores and the total number of phosphopeptides identified is used as an indication of the ability to identify sites of phosphorylation.

| Peptide                                 | CID (MSA) 1 <sup>st</sup> and 2 <sup>nd</sup> ID |                                  | ETD 1 <sup>st</sup> and 2 <sup>nd</sup> ID |                 |
|---|--|----------------------------------|--|-----------------|
|   | 1 <sup>st</sup>                                  | 2 <sup>nd</sup>                  | 1 <sup>st</sup>                            | 2 <sup>nd</sup> |
| IAYSDFETLKVDFLSKLP EMLK+p               | 56   | 33                               | 93   | 41              |
| DFETLKVDFLSKLP EMLK+p                   | 10   | 9 (3 <sup>rd</sup> ID correct)*  | 65   | 20              |
| LKLSSEDSLTKQPEEVFDVLEK+p                | 57   | 55                               | 26   | 24              |
| LKLSSEDSLTKQPEEVFDVLEK+2p               | 53   | 51                               | 55   | 47              |
| KLSEDSLTKQPEEVFDVLEK+p                  | 78   | 45                               | 71   | 69              |
| KLSEDSLTKQPEEVFDVLEK+2p                 | nd   | nd                               | 24   | 24              |
| LSEDSLTKQPEEVFDVLEK+p                   | 69   | 65                               | 65   | 49              |
| LRNKTLEDEIATLK+p                        | nd   | nd                               | 55   | 14              |
| NKTLEDEIATLK+p                          | 74   | 17                               | 48   | 11              |
| LADFGVAGQLTDTMAK+p                      | 80   | 62                               | 111  | 80              |
| LADFGVAGQLTDTMAK+p                      | 100  | 78                               | 66   | 47              |
| ATATQLLQHPFK+p                          | 62   | 59                               | 56   | 46              |
| AKRHEEQRELEEEENSDEDELDSHTMVK+2p         | nd   | nd                               | 43   | 24              |
| RHEEQRELEEEENSDEDELDSHTMVKTSVESVGTMR+2p | nd   | nd                               | 42   | 25              |
| RHEEQRELEEEENSDEDELDSHTMVK+p            | nd   | nd                               | 119  | 51              |
| RHEEQRELEEEENSDEDELDSHTMVK+2p           | nd   | nd                               | 111  | 87              |
| HEEQRELEEEENSDEDELDSHTMVK+p             | nd   | nd                               | 73   | 42              |
| HEEQRELEEEENSDEDELDSHTMVK+2p            | nd   | nd                               | 57   | 41              |
| NKSHENCQNMHEFPFMSK+p                    | 53   | 13                               | 63   | 10              |
| VPQGDGDFLKNLSELELQMR+p                  | 104  | 29                               | 86   | 21              |
| NLSELELQMR+p                            | 30   | 18                               | 51   | 13              |
| ALDPMMERIEELRQRYTAK+p                   | 17   | 15 (6 <sup>th</sup> ID correct)* | 45   | 43              |
| QRYTAKRQPILDAMDAK+p                     | 23   | 19 (2 <sup>nd</sup> ID correct)* | 70   | 57              |
| YTAKRQPILDAMDAK+p                       | nd   | nd                               | 63   | 61              |

1<sup>st</sup>-2<sup>nd</sup>/1<sup>st</sup> and 1<sup>st</sup>-2<sup>nd</sup> CID 0.27 and 19.4 (n=15) ETD 0.37 and 25 (n=24)

## Conclusions

ETD and CID are complementary fragmentation techniques with preferential fragmentation of particular charge states. Of importance to the study of protein phosphorylation is the nature of the peptide fragmentation induced by electron transfer dissociation. Primarily c and z type product ions are generated *without neutral loss*. Thus, for example, phosphothreonine can be detected in an ETD MS/MS spectrum by the observation of pairs of peaks differing in mass by 181. This leads to enhanced interpretability of ETD MS/MS spectra relative to CID. *Work here demonstrates that ETD is better able to discern the location of sites of phosphorylation than is CID (MSA). Additionally, more phosphopeptides were confidently identified with ETD than with CID (MSA).*

## References

Schroeder et al., Anal. Chem. (2004) 76, 3590-3598  
Syka et al., PNAS USA (2004) 101, 9528-9533

© 2007 Thermo Fisher Scientific Inc. All rights reserved. Mascot® is a trademark of Matrix Science. SEQUEST is a registered trademark of the University of Washington. All other trademarks are the property of Thermo Fisher Scientific Inc. and its subsidiaries.