Title: Identifying Unknown Enzyme-Substrate Pairs from the Cellular Milieu with Native Mass Spectrometry

Authors: Kalli C. Catcott, Jing Yan, Wanlu Qu, Vicki H. Wysocki, and Zhaohui Sunny Zhou

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: ChemBioChem 10.1002/cbic.201600634

Link to VoR: http://dx.doi.org/10.1002/cbic.201600634
Identifying Unknown Enzyme-Substrate Pairs from the Cellular Milieu with Native Mass Spectrometry


Abstract: The enzyme-substrate complex is inherently transient, rendering its detection difficult. In our framework, IsoLAIT (Isotope-Labeled, Activity-based Identification and Tracking), designed for bisubstrate systems, the common substrate, such as S-adenosyl-L-methionine (AdoMet) for methyltransferases, is replaced by an analogue (e.g., S-adenosyl-L-vinthionine) that, as a probe, creates a tightly bound [enzyme•substrate-probe] complex upon catalysis by thiopurine-S-methyltransferase (TPMT, EC 2.1.1.67). Then, this persistent complex is identified by native mass spectrometry from the cellular milieu without separation. Furthermore, the probe’s isotope pattern flags even unknown substrates and enzymes. IsoLAIT can be broadly applicable for other enzyme systems, particularly those catalyzing group transfer and with multiple substrates, such as glyceraldehyde-3-phosphate dehydrogenase and kinases.

In the complex cellular milieu, understanding which enzyme catalyzes which reaction is paramount for deciphering biology and disease. There are still many enzymes whose exact functions or substrates remain unknown. Conversely, we know the nature of some biotransformations, but not precisely the responsible enzymes or proteoforms (specific genetic and splicing variations including any post-translational modifications). For example, more than 80 methyltransferases exist in humans; and while histone substrates are well characterized, non-histone substrates remain poorly understood.[2, 3]

Screening is a common approach for identifying enzyme-substrate pairs.[4-6] For enzyme families with common substrates, such as S-adenosyl-L-methionine (AdoMet or SAM) for methyltransferases or adenosine triphosphate (ATP) for kinases, radiolabeled or affinity-labeled substrates, such as biotin-labeled ATP, or biorthogonal-tagged analogues, such as propargyl and ketone AdoMet analogues, have been used to screen for enzyme-substrate pairs.[7, 8] However, this methodology fails to identify specific enzyme-substrate pairs, particularly from cellular contexts as multiple substrates can be tagged by multiple enzymes with no clear path for deconvolution.

Identifying enzyme-substrate pairs is hampered by the fact that their interactions are inherently transient. In a catalytic cycle (Figure 1A), the enzyme binds the substrates, converts to products, and releases them. With few exceptions,[15-18] ternary complexes are generally weakly bound, so cannot be directly observed by mass spectrometry (MS). However, non-catalytic complexes are observed by native MS. Enhancing the enzyme-substrate affinity is one key to successful detection.

Towards this end, we envision a new enzyme-substrate pair detection platform, dubbed IsoLAIT (Isotope-Labeled, Activity-Based Identification and Tracking). IsoLAIT is a method that links solution-phase activity to gas-phase detection by using a probe to capture, and native MS to identify, enzyme-substrate pairs from a single sample run with minimal sample preparation (Figure 1B).

IsoLAIT has three keys: first, a probe that preserves the interaction between the enzyme and substrate for analysis; next, a detection method that retains enzyme-substrate interactions; and finally, a unique isotopic flag that clearly identifies the complex, without a priori knowledge of its constituents.

First, the IsoLAIT probe forms an activity-based bisubstrate adduct (Scheme 1), which has sufficient affinity to the enzyme such that the [enzyme•substrate-probe] complex remains intact throughout subsequent analysis. Additionally, high affinity and low turnover enriches the [enzyme•substrate-probe] complex over the transient catalytic complex.

Second, the complex is analyzed by native MS, which maintains protein complexation in the gas phase. Augmented with nano-electrospray ionization, native MS has emerged as a powerful way to characterize protein-ligand binding.[16-20] Upon dissociation, the enzyme and substrate-probe are separated and can be further interrogated, elucidating the identity and structure of each component. An added benefit, the complex is detected in multiple charge states (Figure 2), providing redundancy in complex systems where peaks overlap and helping to deduce the enzyme mass.

Third, the unique isotopic flag conferred by the probe is used as a telltale-sign for the [enzyme•substrate-probe] complex. Overcoming the limitation of typical workflows that require predefined masses, we are able to identify unknown substrates and enzyme proteoforms with unforeseen modifications.

Herein, the IsoLAIT framework is demonstrated with methyltransferases, a large family of enzymes with diverse substrates. In transmethylation, the enzyme transfers a methyl from common donor AdoMet to a nucleophilic substrate—such as thiopurine methyltransferase[21] (TPMT, EC 2.1.1.67) which methylates thiophenols (Scheme 1A). For this system the selected IsoLAIT probe was S-adenosyl-L-vinthionine (AdoVin).[22]
Figure 1. (A, inset) In the conventional catalytic cycle, an enzyme binds two substrates to create a ternary [enzyme•substrate•substrate] complex. After catalysis, the [enzyme•product•product] complex dissociates, products are released, and the apo-enzyme continues on to substrate binding. The transient enzyme-substrate interactions often do not have sufficient affinity to be analyzed by mass spectrometry. (B) Alternatively, the IsoLAIT platform uses a substrate analogue as a probe. An enzyme catalyzes bisubstrate adduct formation between the probe and the enzyme’s native substrate. The resulting tightly bound [enzyme•substrate-probe] complex can be analyzed by native mass spectrometry. Tandem mass spectrometry dissociates the enzyme from the adduct. The probe’s isotopic flag is easily identified, and serves as a reporter for such complexes; thereby, unknowns can be identified without a priori knowledge of the mass of the components involved. Further analysis can be used to elucidate the structures and sequences of the substrates and enzyme from the complex.

We posited that TPMT would catalyze the formation of a substrate-probe adduct[7, 24-29] (Scheme 1B); indeed adduct formation between the substrate’s thiol and the vinyl sulfonium in AdoVin was confirmed.[22] However, it was not clear whether such an [enzyme-substrate-probe] complex would bind sufficiently to survive native MS. To investigate, we prepared both in vitro and ex vivo (i.e., E. coli cell lysate) samples (Figure S1). AdoVin was prepared enzymatically[22] with S-adenosylmethionine synthetase (MAT, EC 2.5.1.6) using both labeled (+15 Da) and unlabeled ATP along with vinthionine (Figure S1).

For the IsoLAIT platform, minimal sample preparation is required. The reaction mixtures were first exchanged into a volatile buffer near physiological pH (e.g., ammonium acetate adjusted to pH 8.0 with ammonium hydroxide) to maintain native conformations and enhance ionization, and then directly infused into the mass spectrometer.

From the in vitro samples, the [enzyme-substrate-probe] complex and apo-enzyme were readily detected and also resolved from each other (Figure 2A, inset). Tandem MS (e.g., collision induced dissociation (CID) at a collision energy of 500 eV) resulted in the apo-enzyme and the substrate-probe adduct (e.g., TNB-AdoVin) (Figure 2B). The latter was easily spotted from its signature 15 Da doublet imparted by the isotopically labeled AdoVin probe (Figure 2B, inset).

Further fragmentation of the TNB-AdoVin adduct was achieved by using higher collision energies for CID (1200 eV) (Figure S2) or quasi-MS3 analysis (Figure S3). The resulting fragments confirmed the identity of the TNB substrate.

Scheme 1. (A) S-adenosylmethionine (AdoMet or SAM), is used by thiopurine methyltransferase (TPMT) to methylate aromatic thiols and produces S-adenosylhomocysteine (AdoHcy or SAH). (B) S-Adenosylvinthionine (AdoVin), an AdoMet analogue and methyltransferase probe, forms a covalent bond with aromatic thiols under TPMT catalysis and results in a substrate-probe adduct that tightly binds to the enzyme.
For internal use, please do not delete. Submitted_Manuscript

Figure 2. (A) In vitro [TPMT-TNB-AdoVin] complex and apo-TPMT were detected in multiple charge states using native mass spectrometry. In this simple system, the bound and apo-forms of TPMT were easily distinguished (inset). (B) Collision induced dissociation (CID) at a collision energy of 500 eV of the 10+ charged precursor ion of the [TPMT-TNB-AdoVin] complex (*) in (A). This collision resulted in some apo-TPMT with the 10+ and 9+ charge states and the TNB-AdoVin adduct (highlighted and inset). Doublets observed in the TNB-AdoVin spectra are due to a mixture of natural and isotopically labeled AdoVin (+15 Da). (C) [TPMT-AMBA-AdoVin] complex was identified in multiple charge states from whole cells without purification. In this complex matrix, bound and apo-forms of TPMT were not resolved (inset). (D) CID at a collision energy of 500 eV of the 10+ charged precursor ion of the [TPMT-AMBA-AdoVin] complex (*) in (C). After collision, again, multiple charge states for the apo-enzyme were observed along with the AMBA-AdoVin adduct (highlighted and inset). In this case, isotopically labeled probe was not used, resulting in a single peak. Isotopically labeled spectra were previously reported. [22]

As compared to in vitro results, it is not clear whether both adduct-bound and apo-forms of the enzyme from the ex vivo sample were present (Figure 2C, inset). This could be due to peak overlap, a common issue for the analysis of complex biological samples, as well as overall signal strength.

It is worth noting, the mass charge ratio for the ex vivo sample differs from that of the in vitro sample. Tandem MS of the ex vivo samples revealed the substrate-probe adduct to be unexpectedly modified. The nitro group of the TNB substrate was reduced to an amine (2-amino-5-mercaptobenzoic acid or AMBA) in the cell lysate, likely due to endogenous nitroreductases,[22, 30] and the corresponding amine-containing substrate-probe adduct (i.e., AMBA-AdoVin) was formed (Figure 2D). Conventional workflows would likely miss this species as the mass changes were unexpected; however, IsoLAIT is not bound by predefined masses, but screens for mass patterns imparted by the probe and thus will identify substrates no matter their modifications.

Equally interesting is the ability to identify the functional enzyme proteoform. For tandem mass spectra, identifying the apo- and adduct-bound enzyme peaks was straightforward. Upon closer analysis of apo-TPMT, the observed and theoretical masses differed—the enzyme underwent methionine cleavage when expressed in E. coli (Figure S4). These results highlight IsoLAIT’s utility in identifying an unknown, biologically-relevant, active proteoform.

Complex dissociation using surface induced dissociation (SID) instead of CID yielded similar results (Figure S5). Though redundant in the case of the monomeric TPMT enzyme, as interest in characterizing multimers or protein complexes increases,[20] SID’s utility is significant.[31, 32]

Throughout screening, we attempted to identify the enzyme-substrate complex without the use of a probe. Though TPMT was observed in complex with the common substrate (AdoMet) and byproduct S-adenosylhomocysteine (AdoHcy or SAH) (Figures S6, S7), we did not observe the TPMT enzyme bound to either TNB or AMBA substrates without the use of the AdoVin probe, again highlighting the transient nature of substrate-enzyme interactions.

Through the given examples, IsoLAIT demonstrates its robustness as a general platform, succeeding in contexts where conventional workflows may fail. Peak overlap in complex matrices is overcome through using the probe as a flag in the tandem mass spectra. Importantly, this method requires no prior knowledge of the specific enzyme or substrate for identification. The unexpected alteration of substrate observed in the ex vivo samples perfectly illustrate both the challenges associated with biological systems and the utility of our methods. Had the ex
The reaction proceeded in a 50 mM potassium phosphate buffer (pH 8.0), 5 mM KCl, 2.5 mM MgCl₂ and unlabeled ATP (10 mM) along with vinthionine (1 mM). The vinthionine [CAS 83768-89-2] was prepared enzymatically[22] using adenosyl-methionine synthetase (MAT, EC 2.5.1.6) using both labeled and unlabeled substrates.

The IsoLAIT framework solves two key challenges in detecting enzyme-substrate pairs: first, how to overcome the transient nature of enzyme-substrate interactions, and second, how to identify components from a complex mixture with unexpected modifications. The former—transient interaction—is overcome by the combination of a probe that perpetuates enzyme-substrate interactions and native MS that retains binding in the gas phase. The latter—identification—is accomplished by using an activity-based probe and monitoring the telltale isotopic pattern.

This appears to be the first reported example of mass spectrometric identification of enzyme-substrate complexes from the cellular milieu, but it need not be the last. In this case, we used a TPMT specific probe (i.e., AdoVin), but IsoLAIT probes can be tailored for other systems. More promiscuous probes could be used to screen for enzyme-substrate pairs more broadly. For example, aziridinoadenosines have been shown to work as activity-based probes for a range of methyltransferases and form stable, tight-binding bisubstrate adducts.[36, 33-35]

Previous studies using chemical tagging and shotgun proteomics may be amenable for IsoLAIT adaptation.[36, 37] IsoLAIT can be broadly applicable and similarly successful for other enzyme systems, particularly those catalyzing group transfer and with multiple substrates, such as glycosyltransferases and kinases.

**Experimental Section**

**General Procedures** All chemicals with reagent purity or above were purchased from Sigma (St. Louis, MO) and Fisher (Pittsburgh, PA) unless otherwise noted. Immobilized metal ion affinity chromatography (IMAC) was performed on HisTrap HP columns. Ultrafiltration was carried out using filters with 10,000 molecular weight cut off (MWCO).

**Preparation of S-adenosylvinthionine (AdoVin)** S-Adenosyl-L-vinthionine [CAS 83768-89-2] was prepared enzymatically[22] using S-adenosyl-methionine synthetase (MAT, EC 2.5.1.6) using both labeled (+15 Da) and unlabeled ATP (10 mM) along with vinhionine (1 mM). The reaction proceeded in a 50 mM potassium phosphate buffer (pH 8.0), 5 mM KCl, 2.5 mM MgCl₂ and was initiated with 50 mM MAT and incubated at 37 °C. After 2-4 hours incubation, this mixture was used as the in situ probe. AdoVin has similar stability to AdoMet, thus the samples were used immediately and without freezing.

**In Vitro Reactions** For the in vitro samples, His-tagged TPMT was grown in transformed E. coli. Purification and using an IMAC column.[24] Adduct formation was performed in 50 mM potassium phosphate (pH 8.0). The reaction solution contained the in situ AdoVin probe at 1 mM, 500 µM 2-nitro-5-thiobenzoic acid (TNB), 2 mM tris(2-carboxyethyl)phosphine (TCEP), 3.5 µM MTAN and 100 µM TPMT. The reaction was incubated at 37 °C for 3 hours.

**Ex Vivo Reactions** For ex vivo reactions, TPMT-transformed E. coli were grown at 37 °C in LB broth, supplemented with 1 mM isopropl β-D-1-thiogalactopyranoside (IPTG) for 12 hours, then washed three times with lysis buffer (50 mM potassium phosphate, 0.5 M NaCl, pH 8.0). Cells were lysed via sonication on ice, and the cell debris and unbroken cells were removed by centrifugation (8000 x g, 4 °C, and 60 minutes). The supernatant, exogenous reagents were added: in situ AdoVin probe at 1 mM, 500 µM TNB, 2 mM TCEP, 3.5 µM MTAN. The reaction was allowed to react at 37 °C for 3 hours.

**Sample Preparation** Prior to analysis by MS, all samples were buffer exchanged into 20 mM ammonium acetate (adjusted to pH 8.0 with ammonium hydroxide) using at least ten cycles of concentration and dilution in a 10 kDa MWCO centrifuge, ultrafiltration concentrator. Samples were then frozen at -80 °C and thawed immediately prior to analysis. Note: in the case where freezing denatures or unfolds the proteins, samples should be stored at 4 °C or analyzed immediately. Further sample dilution was done in 20 mM ammonium acetate (pH 8). The concentration of ex vivo [TPMT-TNBMethionine] was estimated at 2 µM by SDS-PAGE. The concentration of the in vitro [TPMT-TNBMethionine] was 10 µM as determined by A280nm.

**Native Mass Spectrometry** The nanoelectrospray experiments were performed on a Synapt G2S HDMS (Waters Corp., Wilmslow, UK) with a custom-designed surface-induced dissociation (SID) device installed before the ion mobility cell, as previously described.[30] Each sample was filled into a glass capillary pulled using a Sutter Instruments P-97 micropipette puller (Novato, CA) and electrically connected to high voltage with a platinum wire. The nanoelectrospray source was at a voltage of 1.2-1.5 kV. The sampling cone voltage was set to 20 V and the source offset voltage was set to 20 V to avoid source activation of the complex. Other instrument conditions were 5 x 10⁵ mbar for the source pressure, 2.0 mL/min gas flow rate to the trap cell, 120 µL/min gas flow to the helium cell and 60 mL/min gas flow to the ion mobility cell. The ion mobility wave velocity was 200 m/s and the wave height was 16 V. The TOF analyzer pressure was 1.2 x 10⁶ mbar.

**Tandem Mass Spectrometry** Tandem mass spectrometry experiments were performed via dissociation of the selected ions with collision induced dissociation (CID) and SID. CID experiments were conducted with a trap gas flow rate of 4.0 mL/min and SID was conducted with a trap gas flow rate of 2.0 mL/min. The CID and SID (MS²) experiments were conducted in the trap travelling wave ion guide region before the ion mobility cell. The acceleration voltage in CID and SID was obtained as described previously.[31, 32] The collision energy in eV was calculated via multiplying the acceleration voltage by the charge state of the precursor ion. The fragment ions generated from MS² experiment were separated in the ion mobility cell. The quasi MS² experiment was conducted by selection in the quadrupole, activation in the trap CID cell and separation in the ion mobility cell and activation in the trap travelling wave ion guide.

**Acknowledgements**

We are grateful for the financial support from NSF DBI-0923551 (to V.H.W.) and NSF DBI-1455654 (to V.H.W.) and NIH NIGMS.
IsoLAIT (Isotope-Labeled, Activity-based Identification and Tracking) is a framework designed to identify enzyme-substrate pairs from complex matrices. By replacing a common enzyme substrate with an isotope-labeled, activity-based probe, detection of the [enzyme•substrate-probe] complex by native mass spectrometry is rendered facile, even if the components are of unknown chemical nature and mass.

Kalli C. Catcott, Jing Yan, Wanlu Qu, Vicki H. Wysocki*, Zhaohui Sunny Zhou*

Page No. – Page No.

Identifying Unknown Enzyme-Substrate Pairs from the Cellular Milieu with Native Mass Spectrometry