Published in final edited form as:

Anal Chem. 2016 February 2; 88(3): 1812–1820. doi:10.1021/acs.analchem.5b04098.

UVliPiD: A UVPD-based hierarchical approach for *de novo* characterization of lipid A structures

Lindsay J. Morrison¹, William R. Parker¹, Dustin D. Holden¹, Jeremy C. Henderson², Joseph M. Boll³, M. Stephen Trent², and Jennifer S. Brodbelt^{1,*}

¹Department of Chemistry, University of Texas, Austin, TX 78712

²Department of Infectious Diseases, University of Georgia, Athens, GA 30602

³Department of Molecular Biosciences, University of Texas, Austin, TX 78712

Abstract

The lipid A domain of the endotoxic lipopolysaccharide layer of gram negative bacteria is comprised of a di-glucosamine backbone to which a variable number of variable length fatty acyl chains are anchored. Traditional characterization of these tails and their linkages by nuclear magnetic resonance (NMR) or mass spectrometry is time-consuming and necessitates databases of pre-existing structures for structural assignment. Here, we introduce an automated *de novo* approach for characterization of lipid A structures that is completely database-independent. A hierarchical decision-tree MSⁿ method is used in conjunction with a hybrid activation technique, UVPDCID, to acquire characteristic fragmentation patterns of lipid A variants from a number of Gram-negative bacteria. Structural assignments are derived from integration of key features from three to five spectra and automated interpretation is achieved in minutes without the need for pre-existing information or candidate structures. This utility of this strategy is demonstrated for a mixture of lipid A structures from an enzymatically modified *E. coli* lipid A variant. Twenty-seven lipid A structures were discovered, many of which were isomeric, showcasing the need for a rapid *de novo* approach to lipid A characterization.

Graphical Abstract

Correspondence to: Pavan Muttli. jbrodbelt@cm.utexas.edu; J.S. Brodbelt. Phone (512) 471-0028.

Supporting Information:

Supporting Information includes a more detailed description of the data analysis procedure, a summary of neutral losses and product masses for lipid A molecules, structures of ten lipids used for the training data, and screen-shots of the software format of UVliPiD.

Introduction

The outer leaflet of the outer membrane of Gram-negative bacteria is primarily comprised of a lipopolysaccharide (LPS) layer that is a key virulence factor of Gram-negative pathogens.^{1,2} The extensive structural diversity and impressive complexity of these glycolipid structures makes complete characterization of these molecules challenging. LPS are classified as having three regions: a highly variable O-antigen region, a non-repeating saccharide core, and a hydrophobic, membrane-anchoring domain known as lipid A.³ Lipid A features two $\beta(1,6)$ -linked glucosamine residues, each having multiple fatty acid substituents at the 2, 3, 2' and 3' carbons. The 1 and 4' carbons of the diglucosamine backbone can also be variably decorated with one or two phosphoryl groups, phosphoethanolamine, and glycans.⁴ The lipid A region is primarily responsible for the toxicity of the bacterium; the mammalian Toll-like receptor 4 (TLR4) and MD-2 receptor are sensitive to picomolar levels of lipid A which trigger a signaling cascade leading to production of inflammatory cytokines typically advantageous for clearing the infection.⁵⁻⁷ Overstimulation of the innate immune system, however, can lead to septic shock and even death. One of the structurally demanding aspects of lipid A is that the number and length of the lipid A fatty acid tails and phosphate group modifications vary by species and in response to specific environmental conditions.⁸ Variations in both the phosphorylation pattern and the acyl substituents have additionally been shown to modulate the toxicity of LPS and modified forms of LPS are often antagonists of TLR4. Furthermore, modification of LPS often alters the overall stability of the outer membrane, impacting resistance to antimicrobial compounds.^{3,8-11} Consequently, the development of modified lipid A systems as therapeutics has become an active area of exploration and has increased the need for fast and accurate approaches for analysis of lipid A.¹²⁻¹⁵

Due to the possibility of variable acyl chains, characterization of lipid A poses a unique analytical challenge, most often addressed by tandem mass spectrometry (MS/MS). Measurement of m/z , from which molecular weight can be calculated, by mass spectrometry is typically used in conjunction with NMR to assign or confirm structures based on known structures.¹⁶⁻¹⁸ Although collision induced dissociation (CID) of lipid A typically produces losses of acyl chains, it is difficult from a direct CID approach to assign the location or number of isomeric chains. To overcome this problem, the Goodlett group recently introduced HitMS, a hierarchical tandem mass spectrometry algorithm that uses diagnostic ions and neutral losses from MS^3 and MS^4 fragmentation in conjunction with database

searching to characterize lipid A structures.¹⁹ Correlation scores were used to evaluate the quality of the matches and from this approach 85% of structures from two different databases were correctly assigned.¹⁹ HitMS uses a shotgun approach for MS² and MS³ product ion selection and iteratively interrogates all product ions at each MSⁿ level. The result is the collection of hundreds of spectra, each of which must be parsed prior to structural assignment, a task which reduces the speed of the analysis. Other popular lipid identification algorithms, such as (Analysis of Lipid Experiments, ALEX) are suitable to the analysis of high lipid content lipidomics samples but are not designed to handle the complexity of lipid A.²⁰

The development of other activation methods, such as ultraviolet photodissociation (UVPD),^{21–23} has provided new MS/MS strategies for characterizing structures of biological molecules. Our group has previously shown that ultraviolet photodissociation (UVPD) of lipid A results in the generation of unique glycosidic and cross-ring fragments that can be used to unambiguously identify known lipid A structures at the MS² level.^{10,24–35} Automation of this approach, however, poses problems because the most diagnostic fragment ions are often low abundance and difficult to discern from other non-informative products. Thus, we have designed a streamlined approach that exploits the unique glycosidic cleavages of UVPD. We describe a hierarchical methodology incorporating a new hybrid dissociation technique, UVPDCID, which increases the abundance of diagnostic fragment ions and consequently improves MSⁿ spectral quality. We demonstrate an automated method, UVliPiD, for *de novo* structural characterization of the acyl chain linkages of bis-phosphoryl lipid A structures.

Materials and Methods

E. coli Lipid A was obtained from Avanti Polar Lipids and used without additional purification. Penta and hexa-acyl lipid A species from wild type *Pseudomonas aeruginosa* were isolated and purified as described previously.³² Wild type and lpxL mutant lipid A from *Acinetobacter baumannii* were isolated and purified as described previously.³⁴ Lipid A from *E. coli* expressing the protein LpxJ from *C. jejuni* and *W. succinogenes* was obtained as described previously.³⁰ A modified lipid A was obtained from *E. coli* strain BN2 expressing PagL.¹⁵

Mass spectrometry

Purified lipid A was dissolved in 50:50 chloroform:methanol containing 100 mM ammonium citrate and 0.1% ammonium hydroxide. All MSⁿ experiments were performed on a Thermo Scientific Velos Pro dual linear ion trap mass spectrometer (San Jose, CA) coupled to a 193 nm excimer laser (Coherent, Santa Clara, CA) as described previously.²⁴ Both UVPD and CID can be implemented at different MSⁿ stages. Lipid A molecules were introduced to the gas phase using a static nanospray source comprised of a platinum wire inserted into a pulled glass capillary emitter filled with lipid A in the aforementioned solution. Glass capillaries (1.5 mm OD) were pulled in-house with a Sutter Instrument P-2000 laser puller and had an aperture of less than one micron. Ions were generated by application of 500–1500V to the platinum wire and were detected in the mass spectrometer

under negative ion settings. Ion trap control language (ITCL) coding was modified to allow laser triggering (UVPD) and CID to occur in a two-step manner in the same activation step by incorporation of an isolation toggle. Consequently, UVPD can be applied to an isolated ion and CID can be applied to the precursor or fragment ion without isolation of the second species. In this study, CID was applied the charge-reduced electron photodetachment species (a process known as a-EPD, or activated electron photodetachment dissociation) in all cases. The result of this two-step activation, termed UVPDCID, is the generation of fragments of both UVPD and a-EPD. UVPDCID was performed in all cases using five 2–5 mJ pulses of the excimer laser, pulsed at a frequency of 500 Hz (one pulse every 2 ms). Note that focusing optics were added to the instrumentation for some analyses, and consequently lower laser powers were needed to achieve the same degree of fragmentation. Following UVPD, isolation during the subsequent MSⁿ stage was bypassed and CID was applied to activate the charge-reduced electron detachment species using a normalized collision energy (NCE) of 35. CID for MS² experiments was performed using 25 NCE. CID of the Y₁ and ^{0,4}A₂ ions was performed following UVPDCID or CID-UVPDCID using 23–28 NCE depending on the size of the fragment ion targeted for activation. High resolution data was obtained on a Thermo Scientific Orbitrap Fusion mass spectrometer using the same source and solvent conditions as was used on the Velos Pro mass spectrometer.

Algorithm Development

To simulate a data-dependent acquisition for the *de novo* analysis of Lipid A molecules, a custom automated workflow was designed around a set of rules created based on training data, as described in more detail later. Custom instrument code was developed to implement a decision-tree process and allow automated picking of relevant ions for subsequent isolation and/or activation. These ions were activated to generate key fragment ions and promote diagnostic neutral losses to characterize each lipid A molecule. Next, a stand-alone algorithm was developed in Microsoft Visual Studio 2013 in C# which allows automated characterization based on fragment ions arising from the Y₁, ^{0,4}A₂, X₁, and X₁+52 ions. The cataloging and assignment of the neutral losses of these fragment ions facilitates reconstruction of the lipid A structure. The algorithm is available via this URL: <https://github.com/wrparker/UVliPiD>

Results and Discussion

UVliPiD Strategy

In this work, a hierarchical approach is presented that utilizes UVPDCID, in which a doubly charged (2-), bis-phosphoryl lipid A precursor is subjected to UVPD and the resulting charge-reduced electron photodetachment (EPD) product is activated by CID without an intermediate ion isolation step.³⁶ Bypassing the intermediate ion isolation step is an economical way to implement elaborate MSⁿ schemes and reduces losses of ion current. In order to simplify fragment ion assignments, a nomenclature scheme was developed that encompasses the glycosidic cleavage scheme created by Domon and Costello³⁷ and standard lipid A carbon numbering. This is depicted in Figure 1 for *E. coli* lipid A. Glycosidic fragments are referenced by the nomenclature introduced by Domon and Costello, e.g. A, B, C ions (non-reducing end) and X, Y, Z ions (reducing end). Acyl chain losses are denoted by

the number of the carbon to which the acyl chain is attached and using Greek letters the position of cleavage. Thus, cleavage of the ester bond of the O-linked fatty acid on the reducing sugar is denoted 3β cleavage and cleavage on the glycan side of the amide bond of the N-linked acyl chain of the non-reducing sugar is termed $2'\alpha$ cleavage. A collection of ten lipid A variants from three different species of Gram-negative bacteria were subjected to UVPDCID and the resulting abundances of the fragments resulting from glycosidic cleavages are summarized relative to the abundance of the precursor in Figure 2a. The structures for these ten lipids are shown in Supporting Information Figure S1. Other fragment ion information is also shown in histogram format in Figure 2 based on the hierarchical approach employed in this study (and described in subsequent sections).

Our group has previously shown that two UV-based fragmentation techniques, UVPD and activated electron photodetachment dissociation (a-EPD), generate unique glycosidic fragments of lipid A; these fragments are not generated by CID and can be used to unambiguously assign lipid A structures.²⁵ The instrument control code of a Velos linear ion trap mass spectrometer was modified to allow UVPD and a-EPD processes to be engaged sequentially in a two-step manner in which the second isolation is bypassed. The doubly deprotonated lipid A precursor was activated by UVPD in the first stage of fragmentation and the resulting product ions collected in the ion trap. An abundant population of charge-reduced precursor ions from EPD are generated by UVPD. In contrast to conventional MS³ methods which typically involve isolation and activation of specific product ions, the isolation step was bypassed, and in this case the EPD ions were activated by CID. The acquired MS/MS spectra thus reflect a combination of fragment ions from the two activation methods, and this hybrid activation approach is termed UVPDCID. UVPDCID results in nearly doubling the abundances of many of the fragments compared to a more traditional step-wise MSⁿ method. For example, the Y₁ fragment ion of several lipid A variants was observed in some cases to dramatically increase in relative abundance using the UVPDCID fragmentation approach. Figure 2a shows the relative abundance of several common glycosidic fragments of ten lipid A variants following UVPDCID. Four of the lipid A variants (lipid A from *E. coli* expressing *C. jejuni* LpxJ, lipid A from an *A. baumannii* lpxL mutant, and lipid A from *E. coli* expressing *W. succinogenes* LpxJ) display an enhanced Y₁ ion following UVPDCID, and inspection of the structures of these four variants compared to the six that did not show Y₁ enhancement suggests that lipid A structures lacking a secondary acyl chain at the 2'C position exhibit Y₁ enhancement. From a structural characterization point of view, the generation of this fragment is extremely significant: because bis-phosphoryl lipid A structures have the same phosphate capping group on each termini, it is impossible to differentiate a reducing-end fragment from a non-reducing-end fragment based on mass alone. The generation of a uniquely abundant fragment ion specific to the reducing end of the molecule, however, allows confident designation of the different glycosidic fragments. The increased abundance of the Y₁ ion, moreover, makes it well suited for additional stages of MSⁿ. This reasoning was the motivation for the development of a lipid A *de novo* characterization approach incorporating both directed hierarchical MSⁿ and UVPDCID.

In Figure 2a, six of the ten lipid A variants are observed to generate Y₁ fragment ions, but without the enhanced abundance observed for the four mentioned above (the two lipid A

structures from *E. coli* expressing *C. jejuni* LpxJ, the *A. baumannii* lpxL variant, and the *E. coli* expressing *W. succinogenes* LpxJ lipid A). CID of doubly deprotonated lipid A has been previously shown to predominantly promote acyl losses, and it was postulated that it would be possible to generate the abundant Y_1 ion if a particularly labile acyl chain was first lost from the lipid A structure. The relative abundances (relative to the total ion current) of different fragment ions evolving from acyl chain losses, denoted by the nomenclature shown in Figure 1, are shown in Figure 2b. Some neutral losses of acyl chains result in identical in mass shifts, and the relative abundances of these isomeric losses were determined by ratiometric analysis of the unique glycosidic fragments following subsequent UVPDCID activation of the acyl loss species. Cleavage at the alpha and beta position of primary acyl chains and at the epsilon and zeta position of secondary linkages were considered. The most dominant fragment ions occurred from cleavages at the $3'\epsilon$ or $3'\beta$ positions representing cleavage of the O-linked secondary and O-linked primary acyl chains from the non-reducing sugar, respectively. Cleavage at the $3'\beta$ position, however, was only observed if the lipid A structure lacked a secondary acyl on the $2'C$ position. Consequently, it is possible to confidently predict the most dominant neutral loss upon CID to be the secondary acyl chain, if present, or primary acyl chain linked to the $3'C$ position. If a secondary chain is absent, the primary chain is expected to be lost via cleavage at the beta position. Note that none of the lipid A variants examined herein lacked a primary acyl chain at the $3'C$ position.

In Figure 2c, the dominant neutral loss product created by CID was subjected to UVPDCID (MS^3), and the abundances of the acyl loss and glycosidic fragments were measured relative to the abundance of the precursor. Lipid A variants that generated an abundant Y_1 ion upon UVPDCID (MS^2) were not studied by this CID-UVPDCID approach. All six of the lipid A variants studied produced significant $^{1,5}X_1/Y_1$ pairs by this approach, the abundances of which were three- to four-fold greater than the other fragments. Four of the variants, penta- and hexaacyl lipid A from *P. aeruginosa*, hexa-acyl *A. baumannii*, and lipid A from *E. coli* strain BN2 expressing PagL, also produced dominant $^{1,5}X_1+52$ fragment ions. With the exception of *A. baumannii* hexaacyl lipid A, all of the variants that generated $^{1,5}X_1+52$ fragment ions lacked a secondary acyl chain at the $3'C$ position in the original structure. The relative abundance of the $^{1,5}X_1+52$ product ion from *A. baumannii* hexaacyl lipid A was on average 0.04, only twice the abundance of the “unenhanced” fragment ions (and with a relatively large standard deviation that left room for uncertainty). In sum, CID-UVPDCID generated two or three diagnostic fragments: a $^{1,5}X_1/Y_1$ doublet that allows identification of the reducing end of the lipid A structure and an $^{1,5}X_1+52$ species consistent with lipid A structures altogether lacking acyl chains at the $3'C$ position or having only a primary acyl chain at this position.

A final stage of CID was used to probe the nature of the remaining acyl chains residing on the glycosidic fragments, Y_1 and $^{0,4}A_2$, generated by UVPDCID or CID-UVPDCID, respectively. The fragmentation patterns of these Y_1 or $^{0,4}A_2$ ions are summarized in Figure 2d and 2e, respectively. Activation of the Y_1 ion resulted in exclusive fragmentation at the alpha and epsilon positions (if a secondary chain was present). If a secondary acyl chain was present on the N-linked acyl chain, cleavage at both the 3α and 2ϵ positions was observed, resulting in a doublet approximately 200–300 Da lower in mass than the Y_1 precursor and an additional low abundance fragment ion approximately 500 Da lower than the Y_1 .

corresponding to the loss of both chains. In all cases the 3 α fragment ion was significantly more abundant than the 2 ϵ fragment ion, which allows localization of the different chains based on the fragment abundances. Because none of the lipid A structures contained a secondary acyl chain at the 3C position, it is not clear whether 3 α cleavage would dominate over 3 ϵ cleavage for structures possessing both primary and secondary chains at the 3C position. Fragmentation of the $^{0,4}A_2$ ion was unusual in that both 3' β and 3' α cleavages were observed if an acyl chain was linked to the 3' carbon, resulting in a doublet even if two primary acyl chains were attached to the $^{0,4}A_2$ fragment. The presence of a secondary acyl chain on the N-linked primary chain resulted in the generation of a triplet of peaks approximately 200–300 Da lower in mass than the $^{0,4}A_2$ ion. Structures in which a secondary acyl group was located on the 3'C position resulted in the loss of the secondary and primary acyl chains at the 3' α , 3' β , and 3' ϵ positions, with the fragment specifically evolving from 3' α cleavage being approximately two-thirds the abundance of the 3' β cleavage product and lower in mass by 18 Da. The fragment resulting from 2 ϵ cleavage was the lowest abundance and was approximately 10% the abundance of the 3' β cleavage product, thus allowing distinction of the different chains. Fragmentation of structures having a secondary acyl chain on the 3'C position resulted in cleavage of the secondary acyl chain at the 3' ϵ position as well as cleavage of both chains at the 3' α/β positions. This generates a diagnostic pattern in which there is a single neutral loss of 200–300 Da and a doublet of losses of approximately 500 Da.

Based on the training data from the UVPDCID, CID-UVPDCID, and UVPDCID-CID/CID-UVPDCID-CID experiments described above, an MSⁿ decision-tree workflow for the *de novo* analysis of unknown lipid A structures was developed and is shown in Figure 3. The first step of the analysis involves screening doubly deprotonated lipids by UVPDCID to determine whether or not the lipid A molecules have a secondary acyl chain at the 2'C position. Detection of an abundant singly charged fragment ion indicates the absence of the secondary acyl at the 2'C position. The absence of any dominant peak in this same mass range indicates the presence of a secondary 2'C acyl chain. For lipid A molecules lacking this critical acyl chain, Route 1 of the flowchart is followed in which the abundant singly charged species, which is presumed to be the Y₁ ion, and the companion $^{0,4}A_2$ ion are each subjected to UVPDCID-CID in order to characterize the acyl chains on the reducing and non-reducing glycans. Analysis of lipid A structures that contain a secondary 2'C acyl is inherently more challenging. For these lipids, Route 2 is followed. These lipids are first subjected to CID at the MS² stage in order to remove the primary or secondary (if present) 3'C acyl chain. The most abundant doubly deprotonated species that is generated from CID is then analyzed by UVPDCID. Diagnostic fragment ions 28 Da apart are used to identify the $^{1,5}X_1$ and Y₁ products. The appearance of a fragment 52 Da higher in mass than the X₁ ion confirms that the lipid A structure lacks a secondary 3'C acyl group. The Y₁ ion and presumably truncated $^{0,4}A_2$ ion, which can be determined from the masses of the Y₁ and precursor ions, are then characterized by another stage of CID (CID-UVPDCID-CID) to obtain the acyl chain linkages on each side of the lipid A molecule. The elimination of a secondary acyl chain at the ϵ position results in the loss of the acyl chain terminated by a carboxylic acid. The complementary lipid A structure consequently contains an additional double bond on the primary acyl chain on which the secondary chain was localized that

reduces the mass of the primary acyl chain by 2 Da. A reaction scheme depicting the formation of a double bond is shown in Supporting Information Figure S2. Additional details for the *de novo* interpretation are provided in the Supporting Information.

To showcase the capabilities of this decision-tree strategy for lipid A analysis, the ion selection and activation processes illustrated in Figure 3 were automated via custom instrument control code to enable a data-dependent workflow. Structural assignments were automated by an in-house software package developed in C# in Microsoft Visual Studio 2013. This software package incorporates a CSV list of ions of interest identified by the user and derives the predicted lipid A structure. This software package is open-source and freely available to download at <https://github.com/wrparker/UVliPiD>. Interpretation of the UVPDCID/CID-UVPDCID and UVPDCID-CID/CID-UVPDCID-CID spectra and subsequent *de novo* construction of the lipid A structure is approached in a step-wise manner (as described above), and the process is described in even greater detail in Supplemental Information. The outputs from this software are shown in Supporting Information Figure S3.

Determination of Structures for Unknown Lipids

In order to test UVliPiD, a blinded sample was provided and analyzed by the method detailed above. Upon ESI in the negative mode, two doubly charged lipid species of m/z 1017 and 911 were observed in the ESI mass spectrum, and the former was characterized in detail (the abundance of the latter was too low for full characterization). The UVPDCID, CID, CID-UVPDCID and CID-UVPDCID-CID(Y_1)/UVPDCID-CID($^{0,4}A_2$) spectra acquired for the lipid are shown in Figure 4. UVPDCID (Figure 4a) resulted in a number of singly charged ions in the m/z 500–1350 range, none of which had a relative abundance greater than 1.5%. This indicates the lipid A contains both primary and secondary acyl chains at the 2'C position. CID (Figure 4b) of the ion of m/z 1017 resulted in a dominant neutral loss of 228 Da (producing an ion of m/z 903), corresponding to a 14:0 chain lost from the epsilon position. UVPDCID of the ion of m/z 903 (Figure 4c) resulted in a pair of singly charged fragment ions of m/z 976 and 948, differing in mass by 28 Da which identifies these ions as the $^{1,5}X_1$ and Y_1 ions, respectively. The Y_1 (m/z 948) and companion $^{0,4}A_2$ ions (m/z 916, ($\text{mass}_{\text{truncated precursor}} - \text{mass}_{Y_1} + 60$ Da) were subjected to CID, and the resulting spectra are shown in Figure 4d and 4e, respectively.

The masses of the Y_1 and $^{0,4}A_2$ ions are consistent with both having three acyl chains. Three neutral losses are observed in the CID mass spectrum of the Y_1 ion (Figure 4d), resulting in products of m/z 704, 692, and 684 and corresponding to losses of 244, 256, and 264 Da. The ion of m/z 704 is the most abundant and is consistent with loss of a 14:0(3-OH) from the 3 α position. The less abundant ion of m/z 692 produced upon the loss of 256 Da is consistent with the elimination of a secondary chain from the 2 ϵ position; the mass of this neutral loss (256 Da) is consistent with a 16:0 chain. A low abundance ion of m/z 448 reflects the loss of both of these acyl chains and has a mass consistent with a glucosamine moiety bearing a single 14:1 acyl tail. The loss of 264 Da results in a low abundance product and is inconsistent with acyl chain losses; this loss likely corresponds to a different fragment. Activation of the $^{0,4}A_2$ ion (Figure 4e) results in the formation of four products of m/z 672, 690, 708, and 716 that are consistent with acyl chain neutral losses of 244, 226, 208, and

200 Da, respectively. The two most abundant ions (m/z 708 and 690) differ by 18 Da, suggesting these fragments are the result of cleavage at the $3'\beta/\alpha$ positions. This outcome is consistent with the assignment of the O-linked primary 14:1 chain; the double bond is a consequence of loss of the secondary chain. The product of m/z 716 is consistent with neutral loss of 12:0 fatty acid following cleavage at an alpha or epsilon position. Because of the low abundance of the precursor ion (m/z 916), the S/N ratios of the fragment ions resulting from losses of two acyl chains are very low; nonetheless, a low abundance ion of m/z 508 is detected and is consistent with the loss of a 14:0 acyl from the $3'\beta$ position and loss of 12:0 chain from the $2'\epsilon$ position. The product ion of m/z 508 suggests the N-linked primary chain is also a 14:1 chain. The dehydrogenation site on the O-linked chain strongly suggests the 14:0 lost by CID was a secondary chain linked to this primary chain. Figure 5a shows the proposed structure of the re-constructed lipid A (nominal molecular weight 2034 Da). Interestingly, the ion of m/z 690 is 18 Da greater than the ion of m/z 672 (Figure 4e, CID of $^{0,4}A_2$ ion) and suggests the possibility of an alternative $3'\beta/\alpha$ cleavage that may indicate the presence of a second isomer. However, the low abundance fragments of m/z 734 and 662 do not allow the assembly of a structure consistent with the mass of the $^{0,4}A_2$ ion. The structure of a second isomer, if present, is not clear at this time.

UVliPiD was subsequently applied to a genetic knockout *E. coli* strain to test this *de novo* strategy for an array of lipid A components. A high resolution mass spectrum of the lipids obtained from the *E. coli* sample is shown in Figure 6, and the various doubly deprotonated species characterized by UVliPiD are labelled with the respective mass-to-charge ratios and subscripts to denote the number of isomeric structures identified for each m/z . The UVliPiD MSⁿ approach was implemented on a low resolution ion trap; however, the molecular compositions of the lipid A species assigned by UVliPiD were verified using high resolution spectra. All structural assignments were made with less than 5 ppm mass error. The abundances of some of the ions in Figure 6 were too low to permit the complete MS⁴ method; these ions are denoted with a double dagger (‡) symbol. Based on the decision-tree *de novo* workflow, the structures assigned to the lipids are summarized in Table 1 using a modified form of the shorthand mass spectrometry lipid nomenclature proposed by Liebisch *et al.* and modified and used elsewhere with lipid A molecules.^{38,39} In total, 27 structures were identified and a large number of isomeric lipid A variants were found, particularly for low abundance species. Interestingly, acyl chains with odd numbers of carbons were also frequently found among the isomeric structures discovered. Examples of this are the four isomers of m/z 890.60₁₋₄, which feature substitution of a 13:0 chain for a 14:0 at each of the primary chains. Although acyl chains with odd numbers of carbons are unusual in *E. coli*, modification of the β -barrel of the PagP enzyme, which is highly specific for 16:0 fatty acid chains, has previously been shown to result in the construction of lipid A molecules with C10, C11, C12, C13, and C15 acyl chains.^{40,41} Interestingly, the relative abundance of the acyl loss fragment ions suggests that incorporation of the 13:0 fatty acid chain at the $2'$ position was the most favorable.^{40,41} Identification and characterization of these isomeric structures highlights the utility of a rapid, database-independent lipid A analysis strategy.

Conclusions

A rapid, database-independent approach is demonstrated for the determination of the lengths and linkages of the fatty acyl tails of bis-phosphoryl lipid A. Structural characterization is achieved following acquisition of three to five MSⁿ spectra, and automation of this process is successful for systematic analysis of an array of lipid A molecules in a data-dependent manner. The utility of a *de novo* method of lipid A characterization is demonstrated for a mixture of lipid A variants, many of which feature isomeric structures. By this strategy, twenty-seven unique structures were discovered, twenty-three of which were isomeric with at least one other structure. Integration of this automated decision-tree approach with chromatographic or ion mobility separations has the potential for even greater depth of analysis.

Lipid A and LPS structures feature a number of variable modifications, including modification of the phosphate groups with phosphoethanolamine and highly variable sugar linkages in the core and O-antigen regions (for LPS). Additional effort is needed to develop and modify UVliPiD for these more complex structures. Database-independent characterization of these larger biomolecules, however, is expected to unveil a more diverse catalogue of structures, identification of which may prove useful to serotyping of Gram-negative bacteria.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding from the NIH (R01 GM103655 to JSB) and the Welch Foundation (F-1155 to JSB) is gratefully acknowledged. Also acknowledged are NIH grants AI064184 and AI076322 (to MST), and grant W911NF-12-1-0390 from the Army Research Office (to MST). LJM acknowledges a fellowship from NIH (1K12GM102745).

References

1. Whitfield C, Trent MS. Annual Review of Biochemistry. 2014; 83:99–128.
2. Needham BD, Trent MS. Nature reviews Microbiology. 2013; 11:467–481. [PubMed: 23748343]
3. Takayama K, Qureshi N, Ribi E, Cantrell JL. Reviews of infectious diseases. 1984; 6:439–443. [PubMed: 6474003]
4. Rietschel E, Wollenweber HW, Zähringer U, Lüderitz O. Klin Wochenschr. 1982; 60:705–709. [PubMed: 6750222]
5. Poltorak A, He X, Smirnova I, Liu MY, Van Huffel C, Du X, Birdwell D, Alejos E, Silva M, Galanos C, Freudenberg M, Ricciardi-Castagnoli P, Layton B, Beutler B. Science. 1998; 282:2085–2088. [PubMed: 9851930]
6. Hoshino K, Takeuchi O, Kawai T, Sanjo H, Ogawa T, Takeda Y, Takeda K, Akira S. Journal of immunology (Baltimore, Md: 1950). 1999; 162:3749–3752.
7. Shimazu R, Akashi S, Ogata H, Nagai Y, Fukudome K, Miyake K, Kimoto M. The Journal of experimental medicine. 1999; 189:1777–1782. [PubMed: 10359581]
8. Raetz CR, Reynolds CM, Trent MS, Bishop RE. Annu Rev Biochem. 2007; 76:295–329. [PubMed: 17362200]

9. Post DMB, Phillips NJ, Shao JQ, Entz DD, Gibson BW, Apicella MA. *Infect Immun*. 2002; 70:909–920. [PubMed: 11796626]
10. Herrera CM, Hankins JV, Trent MS. *Mol Microbiol*. 2010; 76:1444–1460. [PubMed: 20384697]
11. Lee C, An H-J, Kim J-I, Lee H, Paik S-G. *Mol Cells*. 2009; 27:251–255. [PubMed: 19277509]
12. Carillo S, Pieretti G, Bedini E, Parrilli M, Lanzetta R, Corsaro MM. *Carbohydr Res*. 2014; 388:73–80. [PubMed: 24632212]
13. Chebrolu C, Artner D, Sigmund AM, Buer J, Zamyatina A, Kirschning CJ. *Mol Immunol*. 2015; 67:636–641. [PubMed: 26319313]
14. Badmasti F, Shahcheraghi F, Siadat SD, Bouzari S, Ajdary S, Amin M, Halabian R, Imani Fooladi AA. *J Mol Microbiol Biotechnol*. 2015; 25:37–44. [PubMed: 25677974]
15. Needham BD, Carroll SM, Giles DK, Georgiou G, Whiteley M, Trent MS. *Proc Natl Acad Sci U S A*. 2013; 110:1464–1469. S1464/1461–S1464/1445. [PubMed: 23297218]
16. Schilling B, McLendon MK, Phillips NJ, Apicella MA, Gibson BW. *Anal Chem*. 2007; 79:1034–1042. [PubMed: 17263332]
17. Casabuono AC, van der Ploeg CA, Roge AD, Bruno SB, Couto AS. *Rapid Commun Mass Spectrom*. 2012; 26:2011–2020. [PubMed: 22847700]
18. Wang Z, Li J, Altman E. *Carbohydr Res*. 2006; 341:2816–2825. [PubMed: 17049500]
19. Ting YS, Shaffer SA, Jones JW, Ng WV, Ernst RK, Goodlett DR. *J Am Soc Mass Spectrom*. 2011; 22:856–866. [PubMed: 21472520]
20. Husen P, Tarasov K, Katafiasz M, Sokol E, Vogt J, Baumgart J, Nitsch R, Ekroos K, Ejlsing CS. *PLoS ONE*. 2013; 8:e79736. [PubMed: 24244551]
21. Brodbelt JS. *Chem Soc Rev*. 2014; 43:2757–2783. [PubMed: 24481009]
22. Ly T, Julian RR. *Angew Chem, Int Ed*. 2009; 48:7130–7137.
23. Reilly JP. *Mass Spectrom Rev*. 2009; 28:425–447. [PubMed: 19241462]
24. Madsen JA, Boutz DR, Brodbelt JS. *Journal of proteome research*. 2010; 9:4205–4214. [PubMed: 20578723]
25. Madsen JA, Cullen TW, Trent MS, Brodbelt JS. *Analytical Chemistry*. 2011; 83:5107–5113. [PubMed: 21595441]
26. O'Brien JP, Needham BD, Henderson JC, Nowicki EM, Trent MS, Brodbelt JS. *Anal Chem (Washington, DC, U S)*. 2014; 86:2138–2145.
27. Hankins JV, Madsen JA, Giles DK, Brodbelt JS, Trent MS. *Proc Natl Acad Sci U S A*. 2012; 109:8722–8727. S8722/8721–S8722/8729. [PubMed: 22589301]
28. Hankins JV, Madsen JA, Giles DK, Childers BM, Klose KE, Brodbelt JS, Trent MS. *Mol Microbiol*. 2011; 81:1313–1329. [PubMed: 21752109]
29. Hankins JV, Madsen JA, Needham BD, Brodbelt JS, Trent MS. *Methods Mol Biol (N Y, NY, U S)*. 2013; 966:239–258.
30. Cullen TW, O'Brien JP, Hendrixson DR, Giles DK, Hobb RI, Thompson SA, Brodbelt JS, Trent MS. *Infect Immun*. 2013; 81:430–440. [PubMed: 23184526]
31. Rubin EJ, O'Brien JP, Ivanov PL, Brodbelt JS, Trent MS. *Mol Microbiol*. 2014; 91:887–899. [PubMed: 24372821]
32. O'Brien JP, Needham BD, Brown DB, Trent MS, Brodbelt JS. *Chem Sci*. 2014; 5:4291–4301. [PubMed: 25386333]
33. Nowicki EM, O'Brien JP, Brodbelt JS, Trent MS. *Mol Microbiol*. 2014; 94:728–741. [PubMed: 25223756]
34. Nowicki EM, O'Brien JP, Brodbelt JS, Trent MS. *Mol Microbiol*. 2015; 97:166–178. [PubMed: 25846400]
35. Boll JM, Tucker AT, Klein DR, Beltran AM, Brodbelt JS, Davies BW, Trent MS. *MBio*. 2015; 6:e00478–00415. [PubMed: 25991684]
36. Gabelica V, Tabarin T, Antoine R, Rosu F, Compagnon I, Broyer M, De Pauw E, Dugourd P. *Analytical Chemistry*. 2006; 78:6564–6572. [PubMed: 16970335]
37. Domon B, Costello C. *Glycoconjugate J*. 1988; 5:397–409.

38. Liebisch G, Vizcaíno JA, Köfeler H, Trötz Müller M, Griffiths WJ, Schmitz G, Spener F, Wakelam MJO. *Journal of Lipid Research*. 2013; 54:1523–1530. [PubMed: 23549332]
39. Bedoux G, Vallee-Rehel K, Kooistra O, Zahringer U, Haras D. *J Mass Spectrom*. 2004; 39:505–513. [PubMed: 15170746]
40. Khan MA, Moktar J, Mott PJ, Vu M, McKie AH, Pinter T, Hof F, Bishop RE. *Biochemistry*. 2010; 49:9046–9057. [PubMed: 20853818]
41. Khan MA, Neale C, Michaux C, Pomes R, Prive GG, Woody RW, Bishop RE. *Biochemistry*. 2007; 46:4565–4579. [PubMed: 17375935]

Figure 1.

Lipid A fragmentation nomenclature for UVliPiD. Acyl chain losses are shown in blue and are denoted by the number of the carbon of the glycan to which they are attached, and a Greek letter is used to designate which bond of the acyl chain is broken. The most proximal bond to the glycan is designated α , and each subsequent distal bond is β , γ , etc. Standard glycosidic cleavages are indicated by the nomenclature introduced by Domon and Costello³⁷ and are highlighted in red.

Figure 2.

Relative fragment abundances following MS²-UVPDCID, MS²-CID, MS²-CID-MS³-UVPDCID, and MS²-CID-MS³-UVPDCID-MS⁴-CID/MS²-UVPDCID-MS³-CID. The doubly charged precursor is activated by UVPDCID in a) and by CID in b). In c), the doubly deprotonated precursor is activated by CID and the dominant neutral loss fragment is further activated by UVPDCID. In d) and e) the Y_1 and truncated $^{0,4}A_2$ products that arise from UVPD (or CID-UVPDCID) are additionally activated by CID.

Figure 3.
Flowchart for MS/MS decision tree for *de novo* characterization of lipid A variants.

Figure 4.

Characteristic UVliPiD MSⁿ spectra acquired for a lipid A with a nominal molecular weight of 2034 Da. a) UVPDCID (MS²) of the 2- precursor ion (*m/z* 1017), b) CID (MS²) of the 2- precursor ion (*m/z* 1017); c) CID-UVPDCID (MS³) of the ion of *m/z* 903 due to acyl chain loss; d) CID-UVPDCID-CID (MS⁴) of truncated ^{0,4}A₂ (*m/z* 916) and e) CID-UVPDCID-CID (MS⁴) of Y₁ (*m/z* 948). The cleavages from CID and CID-UVPDCID are shown in the upper right portion (f) with the colors of the cleavages corresponding to text of the fragments in b) and c). Fragmentation maps of the truncated ^{0,4}A₂ and Y₁ products are illustrated in the lower right (f) such that 3'ε and 3ε cleavages are shown in blue and 2'α/β and 2α cleavages are shown in red. The corresponding fragment ion masses are shown in the same colors in d) and e).

Figure 5.
Solved structure of lipid A corresponding to the ion of m/z 1017 (2-)

Figure 6.

High resolution mass spectrum of a genetic knockdown *E. coli* sample. Identified structures are labeled with the m/z of the lipid A ion and isomeric structures are denoted with subscripts. **Table 2** lists the acyl chain pattern of the structures identified. Ions labelled with an asterisk denote singly charged species. The abundances of the ions labelled with a double dagger were too low to permit complete MS⁴ analysis.

Lipid A variants identified from an *E. coli* genetic knockout mutant shown in Figure 6. The m/z labels refer to the monoisotopic m/z of each structure and subscripts denote unique isomeric structures. The acyl chains are denoted by an IUPAC shorthand in which the acyls are listed in order from the 3' carbon of the non-reducing glucosamine towards the 2 carbon of the reducing glucosamine (e.g. 3', 2', 3, 2).

Table 1

m/z	M _{theoretical}	M _{experimental}	C-3'	C-2'	C-3	C-2
679.408	1360.828	1360.831	H	14:0(3-O(12:0))	14:0(3-OH)	14:0(3-OH)
770.492 _{reducing-1}	1542.995	1542.999	---	---	14:0(3-OH)	14:0(3-OH)
770.492 _{reducing-2}	1542.995	1542.999	---	---	13:0(3-OH)	15:0(3-OH)
770.492 _{reducing-3}	1542.995	1542.999	---	---	12:0(3-OH)	16:0(3-OH)
770.492 _{non-reducing-1}	1542.995	1542.999	14:0	14:0(3-O(10:0))	---	---
770.492 _{non-reducing-2}	1542.995	1542.999	14:0	12:0(3-O(12:0))	---	---
777.495 ₁	1557.010	1557.006	14:0(3-O(14:0))	15:0(3-O(10:0))	H	14:0(3-OH)
776.502 ₂	1557.010	1557.006	13:0(3-O(14:0))	15:0(3-O(11:0))	H	14:0(3-OH)
776.502 ₃	1557.010	1557.006	14:0(3-O(14:0))	14:0(3-O(12:0))	H	13:0(3-OH)
776.502 ₄	1557.010	1557.006	13:0(3-O(14:0))	14:0(3-O(13:0))	H	13:0(3-OH)
784.509	1571.026	1571.033	14:0(3-O(14:0))	14:0(3-O(12:0))	H	14:0(3-OH)
792.507	1587.021	1587.029	14:0(3-OH)	14:0(3-O(12:0))	14:0(3-OH)	14:0(3-OH)
806.522	1615.052	1615.060	14:0(3-OH)	14:0(3-O(14:0))	14:0(3-OH)	14:0(3-OH)
876.583 ₁	1755.172	1755.181	14:0(3-O(12:0))	13:0(3-O(12:0))	14:0(3-OH)	14:0(3-OH)
876.583 ₂	1755.172	1755.181	12:0(3-O(15:0))	12:0(3-O(12:0))	14:0(3-OH)	14:0(3-OH)
876.583 ₃	1755.172	1755.181	14:0(3-O(12:0))	14:0(3-O(12:0))	14:0(3-OH)	13:0(3-OH)
876.583 ₄	1755.172	1755.181	14:0(3-O(12:0))	14:0(3-O(12:0))	13:0(3-OH)	14:0(3-OH)
876.583 ₅	1755.172	1755.181	14:0(3-O(12:0))	14:0(3-O(12:0))	15:0(3-OH)	12:0(3-OH)
883.591	1769.188	1769.197	14:0(3-O(12:0))	14:0(3-O(12:0))	14:0(3-OH)	14:0(3-OH)
890.598 ₁	1783.204	1783.212	13:0(3-O(14:0))	14:0(3-O(12:0))	14:0(3-OH)	14:0(3-OH)
890.598 ₂	1783.204	1783.212	14:0(3-O(14:0))	13:0(3-O(12:0))	14:0(3-OH)	14:0(3-OH)
890.598 ₃	1783.204	1783.212	14:0(3-O(14:0))	14:0(3-O(12:0))	13:0(3-OH)	14:0(3-OH)
890.598 ₄	1783.204	1783.212	14:0(3-O(14:0))	14:0(3-O(12:0))	14:0(3-OH)	13:0(3-OH)

m/z	M _{theoretical}	M _{experimental}	C-3'	C-2'	C-3	C-2
904.614 ₁	1811.235	1811.244	14:0(3-O(14:0))	14:0(3-O(13:0))	14:0(3-OH)	14:0(3-OH)
904.61 ₂	1811.235	1811.244	14:0(3-O(15:0))	14:0(3-O(12:0))	14:0(3-OH)	14:0(3-OH)
904.61 ₃	1811.235	1811.244	12:0(3-O(14:0))	16:0(3-O(12:0))	13:0(3-OH)	16:0(3-OH)
904.61 ₄	1811.235	1811.244	12:0(3-O(14:0))	16:0(3-O(12:0))	14:0(3-OH)	15:0(3-OH)
904.61 ₅	1811.235	1811.244	12:0(3-O(14:0))	16:0(3-O(12:0))	15:0(3-OH)	14:0(3-OH)