Check for updates

View Article Online
View Journal

Analyst

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: V. Erckes, A. Streuli, A. Misiek and C. Steuer, *Analyst*, 2026, DOI: 10.1039/D6AN00162A.

This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/analyst

View Article Online
DOI: 10.1039/D6AN00162A

Structure-aware fragment assignment for interpreting tandem mass spectrometry of modified and cyclic peptides

Erckes, V., Misiek, A; Streuli, A.; Steuer, C.*

ETH Zurich, Institute of Pharmaceutical Sciences, Laboratory of Pharmaceutical Analytics, Zurich, Switzerland

* Corresponding author: christian.steuer@pharma.ethz.ch

ORCID

VE: 0000-0002-9650-4160

AS: 0000-0002-0025-8023

CS: 0000-0002-6102-3367

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

Abstract

Peptidomimetics such as fatty-acid-modified, head-to-tail cyclic, and disulfide-constrained peptides challenge conventional MS/MS interpretation because their fragmentation pathways extend beyond canonical backbone cleavages or require multiple bond breakages to generate sequence-informative ions. We present a structure-aware algorithm that calculates and labels theoretical fragment ions across modified and cyclic peptides, including side-chain and disulfide-related fragments. To evaluate assignments, we integrate the three numerical metrics sequence coverage, intensity coverage, and signal coverage and assess their behavior across *m/z* tolerances, intensity thresholds, and charge states. Using representative MS/MS data of angiotensin related peptides, liraglutide, semaglutide, cyclosporine, oxytocin, and somatostatin, we demonstrate that the metrics reliably distinguish correct from incorrect assignments, including closely related sequences. Incorporating fatty-acid-specific fragments increased intensity coverage for liraglutide and semaglutide. MS^3 improved sequence coverage for the cyclic peptide cyclosporine relative to MS^2, but no additional benefit was seen at MS^4 level. For disulfide-bonded peptides, the combination of electron-transfer dissociation (ETD) and collision induced dissociation (CID) shifted fragment distributions toward disulfide-cleavage products relative to CID, which were captured by the algorithm's dedicated disulfide labels. Together, these results demonstrate that structure-aware fragment calculation coupled to explicit assignment metrics enables more comprehensive and standardized interpretation of peptidomimetic MS/MS data. With this approach we

View Article Online
DOI: 10.1039/D6AN00162A

lay a foundation for reproducible benchmarking which can be extended to additional cyclicization, modification, and fragmentation techniques.

Introduction

Peptides are an increasingly important class of therapeutic agents in modern drug discovery.[1] While once considered suboptimal drug candidates due to rapid degradation and limited bioavailability, advances in chemical modification have significantly improved their clinical potential.[2-4] Several peptide drugs have successfully reached clinical studies or the market. Examples are semaglutide, which uses a fatty acid conjugation to enhance half-life through albumin binding, or LUNA18, a cyclic peptide with improved membrane permeability and protease stability.[5] These examples highlight how structural modifications, ranging from single amino acid substitutions to cyclization and conjugation, can profoundly modulate pharmacological properties by altering protease stability, target affinity, or cell permeability.[6, 7]

As sequence and structural modifications directly influence peptide drug properties, reliable identification of peptide and peptidomimetic sequences is essential. Tandem mass spectrometry (MS/MS) is the central technique for peptide and protein sequencing because of its high sensitivity, compatibility with complex mixtures, and minimal sample requirements.[8, 9] Alternative approaches such as Edman degradation, nuclear magnetic resonance spectroscopy, and X-ray crystallography provide complementary insights but require large amounts of material and are less suited for high-throughput analysis.[10-14] By comparison, MS/MS enables rapid and automated analysis for the identification of peptides. Peptides are ionized and fragmented, and the resulting fragment ions provide information that can be used to confirm or deduce peptide sequences.[15] Fragmentation methods, such as collision-induced dissociation (CID), higher-energy collisional dissociation (HCD), and electron transfer dissociation (ETD), generate complementary fragmentation patterns that can be exploited to improve sequence coverage.[16-20] However, MS/MS data evaluation generates spectra with hundreds of fragment signals, making manual interpretation time-consuming and creating the need for automated analysis tools.[21]

For linear peptides composed of the 20 canonical amino acids, numerous computational approaches exist to for MS/MS spectra interpretation. Database search algorithms and tools such as SEQUEST[22], Mascot[23], X!Tandem[24], MaxQuant[25] and MS-GF+[26] have been widely adopted in proteomics, enabling rapid peptide identification by matching experimental spectra to theoretical spectra derived from sequence databases.[27] In parallel, de novo sequencing tools including PEAKS[28], Novor[29], DeepNovo[30], and InstaNovo[31] allow direct sequence interpretation without database dependence, facilitating discovery of novel peptides or unexpected sequence variants. However, these methods generally assume linear peptides with canonical residues and provide limited support for chemical modifications or

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

structural constraints. Even when post-translational modifications are considered, they are typically modelled as static mass shifts rather than as distinct fragmentation behaviors. To our knowledge, no comprehensive public spectral database with annotated MS/MS data dedicated to cyclic and backbone-modified peptides is currently available.

Peptidomimetics pose additional challenges. Structural features such as fatty acid conjugation, head-to-tail cyclization, and disulfide bridges create alternative fragmentation pathways, resulting in complex spectra not captured by conventional algorithms. For example, fatty acid-modified peptides such as semaglutide contains additional amide linkages within the modification that might influence fragmentation and complicate spectrum interpretation. Cyclic peptides require multiple fragmentation events to generate sequence-informative ions, and the number of possible fragments increases substantially with peptide length. Multi-stage MS^n has been described to provide this information, but also greatly expands the number of fragment ions and spectra to evaluate. Recent work has demonstrated selective cleavage and linearization of thioether-macrocyclized peptides to facilitate MS/MS interpretation, although this requires prior derivatization of the thioether ring-closing moiety.[32] Disulfide-bridged peptides additionally exhibit fragmentation patterns that depend on whether the disulfide bond remains intact or is cleaved, alongside backbone amide bond cleavage.[33, 34] While dedicated tools exist for specific tasks (e.g., MassMatrix for disulfide mapping[35], or Byonic[36] and pGlyco[37] glycopeptide-focused software for glycans), there is no general framework that systematically accounts for the diverse fragmentation behaviour of modified and cyclic peptides. As a result, interpretation of MS/MS spectra of peptidomimetics is still largely manual, time-consuming, and not standardized.

Recently, we introduced PICKAPEP[38] as a versatile platform for virtual generation and parameter calculation of libraries of modified and cyclic peptides, supporting user-defined residues and diverse cyclization strategies. While valuable for in silico peptide design, its utility for MS/MS experiments remained limited. In this work, we extend PICKAPEP with an algorithm for the calculation of fragment ions characteristic of peptidomimetics. The approach goes beyond canonical backbone fragmentation by considering structural features such as fatty acid conjugation, head-to-tail cyclization and disulfide linkages. We demonstrate its utility using representative MS/MS data from angiotensin peptides, liraglutide, semaglutide, cyclosporine, oxytocin, and somatostatin. Our analysis spans from overall performance across peptide classes to detailed fragmentation behaviour in selected case studies. By enabling structure-aware fragment calculation, the framework supports a more comprehensive and standardized interpretation of MS/MS spectra of peptidomimetics.

Materials and methods

Fragmentation calculations and data processing

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

The fragmentation algorithm was implemented in Python[39] (3.10.10), using the PICKAPEP framework for peptide file generation. Fragmentation reactions were set up with RDKit[40] (2024.03.5). The algorithm supports fragment calculation for linear peptides, peptides with fatty acid side-chain modifications, head-to-tail cyclic peptides, and disulfide-constrained peptides. For each peptide class, theoretical fragment ions are generated up to the precursor charge state, enabling systematic annotation of MS/MS spectra. In addition, the algorithm calculates metrics for fragment assignment of MS/MS data, including sequence coverage, intensity coverage, and signal coverage (Eq. 1-3), providing numerical descriptors of the assignment quality. Metrics are calculated for all signals above a defined intensity threshold within a defined mass-to-charge ratio (*m/z*) tolerance for assignment. Sequence coverage accounts for the type of fragments assigned independently of the individual charge states found.

$$Sequence\ coverage\ (\%) = \frac{number\ of\ type\ of\ fragments\ found}{number\ of\ type\ of\ fragments\ calculated} \times 100\% \qquad (1)$$

$$Intensity\ coverage\ (\%) = \frac{\sum(intensity\ of\ assigned\ peaks)}{\sum(intensity\ of\ all\ peaks)} \times 100\% \qquad (2)$$

$$Signal\ coverage(\%) = \frac{number\ of\ signals\ assigned}{number\ of\ all\ signals} \times 100\% \qquad (3)$$

Data processing and statistical evaluation were performed with NumPy[41] (1.23.5), pandas[42] (1.5.3) and SciPy[43] (1.15.3). Data visualization was carried out with matplotlib[44] (3.10.1) and seaborn[45] (0.13.2). Chemical structures were drawn using ChemDraw (20.0.0.41). Graphical abstract was created in BioRender. Steuer, C. (2026) https://BioRender.com/qlu1rqo.

Peptides

The peptides analyzed in this study were obtained from various sources. The sequences of angiotensin 1-4 (AT1-4), corresponding to the natural angiotensin peptides, were synthesized by automated solid phase peptide synthesis as described in the supporting information (S1). Cyclosporin (CSA) was kindly provided by the Cantonal Hospital of Aarau, Switzerland. The fatty acid-modified peptides liraglutide (LGL) and semaglutide (SGL) were obtained from the commercial formulations Saxenda® (18 mg LGL/3mL LGL, Lot. NP5H973, Novo Nordisk) and Ozempic® dual dose solution for injection (2 mg SGL/1.5 mL SGL, Lot. NP5H991, Novo Nordisk), respectively, and lyophilized prior to use. Somatostatin (SMT, ≥97% HPLC grade) was purchased from Sigma-Aldrich (Buchs, Switzerland).

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D5AN00162A

Oxytocin (OXY) was obtained as a European Pharmacopeia Reference Standard (CRS, Strasbourg, France). Peptides were dissolved at 0.1 mg/mL in water except CSA was dissolved in 10% methanol (MeOH). An overview of all peptides is provided in Table 1.

Tandem mass spectrometry (MS/MS) analysis

ACN and MeOH (OPTIMA, LC-MS grade) were obtained from Fisher Chemicals (Loughborough, UK). Nanopure water was generated using an in-house ELGA Purelab water purification system (VWS, Villmergen, Switzerland). Formic acid (FA; 98.0-100%) was purchased from Sigma Aldrich (Buchs, Switzerland). Peptides were analyzed by reversed phase liquid chromatography coupled to electrospray ionization tandem mass spectrometry (RPLC-ESI-MS/MS). Analyses were performed on a Waters Acquity UPLC system with a C18 Zorbax column (2.1 × 50 mm, 1.8 µm; Agilent, USA), coupled to a heated ESI source and an LTQ XL ion trap mass spectrometer (Thermo Scientific, USA). RPLC analysis was conducted using a binary gradient of solvent A (water + 0.1% FA) and solvent B (ACN + 0.1% FA) with a flow rate of 0.5 mL/min at 25 °C. The gradient was programmed as follows: 0-2 min, 95% A; 2-12 min, linear from 95% to 10% A; 12-15 min, 10% A; 15-20 min, re-equilibration at 95% A. Injection volumes were 2 µL for all peptide samples, except CSA (10 µL). The ESI source was operated with the following parameters: spray voltage 5 kV, capillary temperature 275 °C, capillary voltage 31 V, tube lens 80 V, heater temperature 0 °C (corresponding to approx. 50 °C operating temperature), sheath gas 34 arb, auxiliary gas 11 arb, and sweep gas 0 arb. Full MS scans were acquired in positive ion mode over *m/z* 100-2000 at normal scan rate with data type profile and an AGC target of 1.50 e+04. Data dependent MS/MS with CID using helium 6.0 (Linde Gas, Dagmersellen, Switzerland) as collision gas was performed with an AGC target of 5.00 e+03, a normalized collision energy of 35, an activation Q of 0.25 and an activation time of 30 ms. The most intense precursor ions above a threshold of 3000 counts were selected for fragmentation within *m/z* 250-2000 using a 1.0 *m/z* isolation window. Dynamic exclusion was enabled with the following parameters: repeat count 2, repeat duration 15 s, exclusion list size 50, exclusion duration 15 s, and exclusion mass width ±0.5 *m/z*. Fragment spectra were acquired in centroid mode over *m/z* 110-2000 with a default precursor charge of 3, whereby the lower cutoff was automatically adjusted by the instrument based on precursor mass. Higher-order experiments were included for selected peptides. MS^3 and MS^4 analysis were performed for CSA with the same parameters for CID fragmentation as MS/MS but with a precursor threshold of 100 counts. For SMT and OXY MS^3 experiments with ETD followed by CID were conducted. ETD was performed using fluoranthene (*m/z* 202) as the reagent ion, with a minimum intensity of > 1 x 10^6 in negative mode. ETD operating parameters were emission current 50.00 µA, electron energy -80 V, CI gas pressure 30 psi, source temperature 160 °C, vial 1 temperature 108 °C, restrictor temperature 160 °C and transfer line temperature 160 °C. ETD activation was performed with an activation time of 100 ms and an AGC target of 1.00 e+05. All mass spectrometry data was extracted using XCalibur (Version 4.4.16.14, Thermo Fisher Scientific).

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

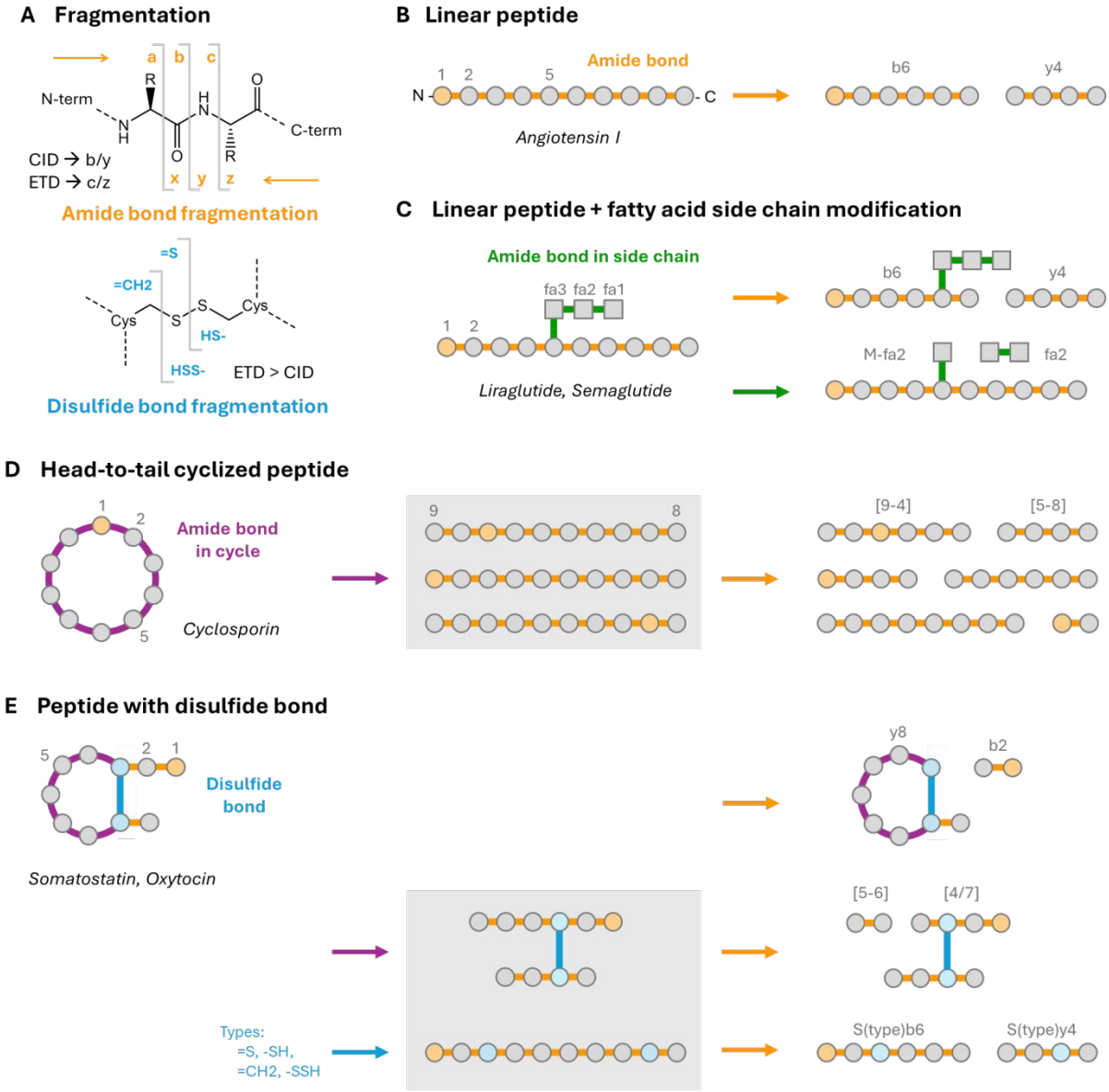
View Article Online
DOI: 10.1039/D6AN00162A

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Results and discussion

Fragment calculation

Efficient evaluation of tandem mass spectrometry data requires both rapid fragment calculation and a consistent nomenclature before data assignment. In this study, fragment generation and calculation was based on PICKAPEP output, which represents peptides through unambiguous numerical encoding of individual atoms in amino acids and modifications.[46] An overview of the principle for fragment generation for the different peptide types is shown in Figure 1. Fragmentation pathways are presented in accordance with reports in literature. Fragments were initially defined in the singly charged state, and their masses were subsequently calculated for higher protonation states.



Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

Figure 1: Fragmentation strategies used for theoretical fragment calculation. (A) amide and disulfide bond fragmentation pathways. Amide bond cleavage predominantly produces b/y-type ions under CID and c/z-type ions under ETD.[16] Disulfide bond cleavage is more common under ETD conditions.[47] (B-E) Principles of fragment generation for the different peptide classes: (B) linear peptides, (C) fatty acid-modified peptides, (D) head-to-tail cyclized peptides, and (E) disulfide-constrained peptides. Fragmentation pathways yielding ions with the same mass as the precursor peptide, which require an additional fragmentation step for sequence information, are highlighted in grey.

Each amide bond can cleave at three different positions, producing fragment ion pairs of the a/x, b/y, or c/z type, corresponding to the N- and C-terminal sides of the peptide, respectively (Figure 1A).[15] Although peptide fragmentation is sequence- and method-dependent, certain method-specific patterns are well established. CID predominantly yields b/y ions, whereas ETD favors c/z ions.[16] Since the final fragmentation step in this study was always performed with CID, calculation and subsequent analysis of linear peptides focuses on b/y-type fragments. In line with established notation, linear peptide fragments were calculated and labelled by ion type and charge state in number of '+' symbols (Figure 1B). For instance, a six amino acid N-terminal fragment detected at a double charge is denoted b6++.

This notation was extended to peptides carrying fatty acid side-chain modifications. To our knowledge, no specific nomenclature has been reported for such cases. In these peptides, amide bonds occur both along the peptide backbone and within the fatty acid-linker connection. Backbone amide cleavages were annotated as in linear peptides, while side-chain cleavages were labelled according to the number of linker segments cleaved, analogous to amino acid numbering. The complementary ion was denoted as the intact molecular ion mass (M) minus the fragmented fatty acid portion (Figure 1C). For example, in liraglutide, which contains a C8 fatty acid (fa1) linked via a γ-glutamate spacer (fa2) to a lysine side chain, cleavage between lysine and glutamate generates two fragments with the notation: fa2 and (M-fa2).

Sequencing head-to-tail cyclized peptides requires at least two cleavage events, resulting in n × (n-1) theoretical b-type fragments for a cyclic peptide with n amino acids and n amide bonds. As described previously, we restrict interpretation to b-type ions while y-type fragments can be omitted. Although alternative ion types (e.g. c/z, a/x) cannot be distinguished based on mass alone in cyclic peptides, this reflects established CID fragmentation preferences.[41] In PICKAPEP, cyclic peptides are created by cyclizing a linear precursor, with residues numbered according to the linear sequence. Fragments were therefore labelled by the first and last amino acid contained in the fragment (Figure 1D). For instance, [5-8] represents cleavage between residues 4/5 and 8/9, corresponding to the sequence from residue 5 to 8. Numbering can extend across the cycle, as for example [9-4] denotes the complementary segment of [5-8].

For peptides containing side chain-to-side chain disulfide bridges, three types of fragmentation pathways were considered (Figure 1E). Backbone cleavages of amide bonds outside the cyclic region

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

were annotated as in linear peptides. Cleavages within the cyclic region produced either continuous fragments labelled as in head-to-tail cycles, e.g. [5-6] or discontinuous fragments, labelled by the two residues retained in the cycle, e.g. [4/7]. Disulfide bond cleavages were previously described to occur either at the S-S bond or the C-S bond, yielding to eight distinct fragmentation products per disulfide bridge. These fragments were annotated as linear ions, with additional labels denoting the disulfide modification state (e.g. =S, -SH, $=CH_2$, -SSH) resulting in for example a fragment S(=S)b6. The opening of the disulfide bond in MS/MS analysis has especially been described in ETD.[47]

Evaluation of assignment metrics for peptide discrimination

An overview of all peptides used in this study with their calculated *m/z* values is provided in Table 1. For all peptides the LC-MS and normalized MS/MS spectra recorded can be found in the supporting information (S2). For each peptide the observed charge states are highlighted in Table 1. For evaluation, only fragment spectra derived from the monoisotopic precursor were used, excluding contributions from isotopic precursor distributions. The $[M+H]^+$ species of LGL and the $[M+H]^+$ and $[M+2H]^{2+}$ species of SGL exceeded *m/z* 2000 and could not be analyzed with our settings.

All spectra were assigned using the developed algorithm, and sequence coverage, intensity coverage and signal coverage were calculated for each assignment. An example output for AT1 $[M+3H]^{3+}$ is shown in Figure 2. The assigned spectra of all MS/MS data can be found in the supporting information (S3).

Table 1: Overview of peptides with underlying amino acid sequence and theoretical *m/z* values expected for ESI MS analysis were calculated using PICKAPEP.[38] *M/z* values found in MS analysis are highlighted in black, not detected *m/z* values are shown in grey. Values in italic could not be detected because their values exceeded the upper limit of the detection range with a *m/z* of 2000.

Peptide	Type	Sequence	$[M+H]^+$	$[M+2H]^{2+}$	$[M+3H]^{3+}$	$[M+4H]^{4+}$
AT1	Linear	H_2N-DRVYIHPFHL-COOH	1296.69	648.85	432.90	324.93
AT2	Linear	H_2N-DRVYIHPF-COOH	1046.54	523.78	349.52	262.39
AT3	Linear	H_2N-RVYIHPF-COOH	931.52	466.26	311.18	233.63
AT4	Linear	H_2N-VYIHPF-COOH	775.41	388.21	259.14	194.61
LGL	Linear + fatty acid	H_2N-HAEGTFTSDVSSYLEG QAAK(*)EFIAWLVRGRG-COOH *γ-Glu-C16 fatty acid	*3749.95*	1874.98	1250.32	937.99
SGL	Linear + fatty acid	H_2N-HAibEGTFTSDVSSYLEG QAAK(*)EFIAWLVRGRG-COOH *2 Ado-γ-Glu-C18 fatty diacid	*4112.12*	*2056.06*	1371.04	1028.53

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D5AN00162A

CSA	Cyclic (head-to-tail)	V(1)-Lnme-A-Ad-Lnme-Lnme-Vnme-Mebmt-Abu-Sar-Lnme(1)	1202.85	601.42	401.29	301.22
OXY	Cyclic (disulfide)	H_2N-C(1)YIQNC(1)PLG-Am	1007.44	503.72	336.15	252.37
SMT	Cyclic (disulfide)	H_2N-AGC(1)KNFFWKTFTSC(1)	1637.72	818.86	546.24	409.94

Ado = 3,8-dioxaamino octanoic acid, Abu = Aminobutyric acid, Sar = Sarcosine, -nme = N-methylated, -d = D-amino acid

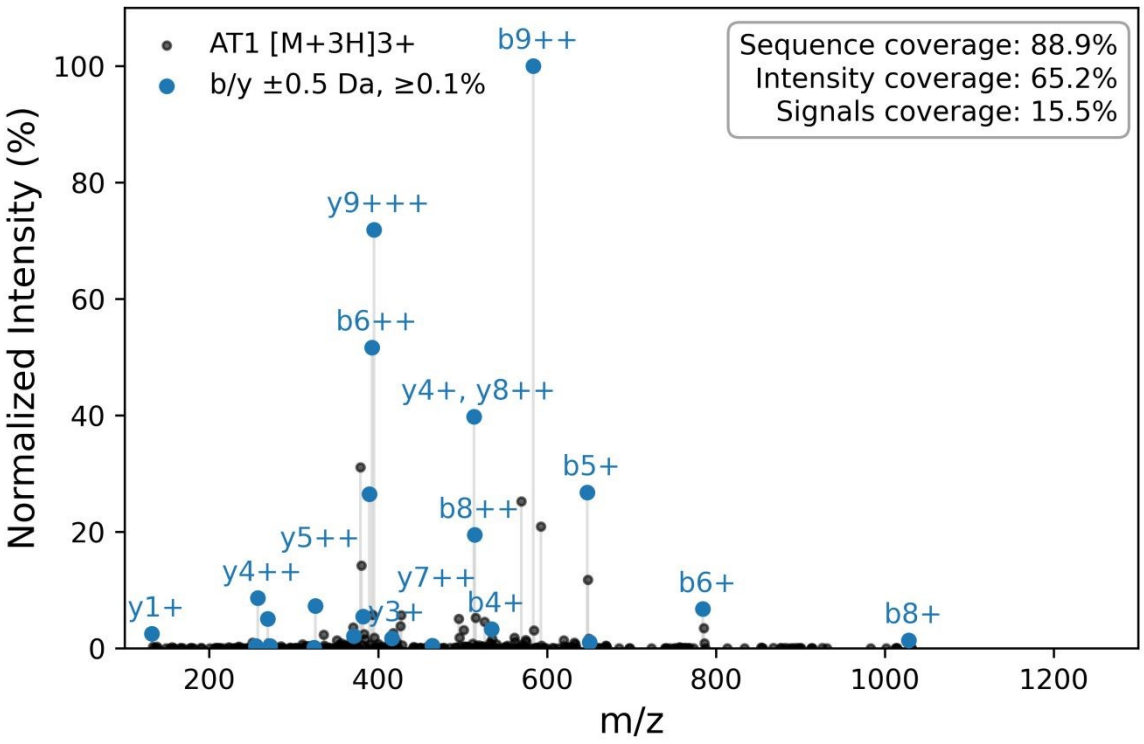


Figure 2: Example of the output for our algorithm-based MS/MS data assignment for the spectrum of AT1 $[M+3H]^{3+}$. All assigned fragment ions are shown in blue, with only a subset labelled for clarity.

We first assessed the applicability of the metrics to linear peptides, as their fragmentation behaviour is well characterized in the literature and directly aligns with the fragmentation pathways implemented in our algorithm.[15, 16] Using AT1 $[M+3H]^{3+}$ as an example (Figure 2), we assessed the sensitivity of assignment metric values to *m/z* tolerance and intensity threshold. As shown in Figure 3A, sequence coverage increased with broader *m/z* tolerances and lower thresholds. Intensity coverage increased with *m/z* tolerance until reaching a plateau and showed little dependence on the intensity threshold, reflecting the dominance of high-intensity peaks in this metric. Signal coverage behaved differently, increasing with both tolerance and threshold, suggesting higher-intensity peaks are more likely to relate to the known fragment patterns compared to lower intensity fragments. Because coverage metrics remained stable at higher *m/z* tolerances and considering the resolution of our instrument, a tolerance of ±0.5 *m/z* was selected for subsequent analyses unless stated otherwise. With low-resolution MS, some degree of

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

unspecific fragment assignment cannot be fully excluded. However, we never observed signals above the precursor mass in any spectrum, even when scans extended beyond this range due to default charge state settings. This suggests that the large number of low-intensity signals does not represent random background noise. Instead, they may arise from sequence scrambling or other poorly understood fragmentation pathways, and we therefore interpret them as potential but unspecific fragment assignments. Instead, they may arise from sequence scrambling or other poorly understood fragmentation pathways, and we therefore interpret them as potential but unspecific fragment assignments.

Based on these considerations, we also wanted to evaluate the ability of the metrics to detect false assignments. For a first evaluation, the AT1 $[M+3H]^{3+}$ spectrum was compared with theoretical fragments from AT2 (related sequence with theoretical fragment overlap of 50%) and CSA (unrelated but with >100 theoretical fragments) as shown in Figure 3B. Heatmaps and threshold-dependent plots of other charged states for the comparison of AT1, AT2 and CSA are provided in the Supporting information (S4). As expected, AT1 values were closer to those of AT2 than to CSA, while overall the metrics showed lower values for incorrect assignments, particularly at higher intensity thresholds. The comparison with CSA also showed a measurable level of matching despite the absence of sequence relatedness. Given the large number of theoretical fragments for CSA, particularly when multiple charge states are considered, this can be attributed to an increased probability of incidental matches within the applied m/z tolerance. This therefore reflects a baseline level of unspecific matching that should be considered when interpreting assignment metrics.

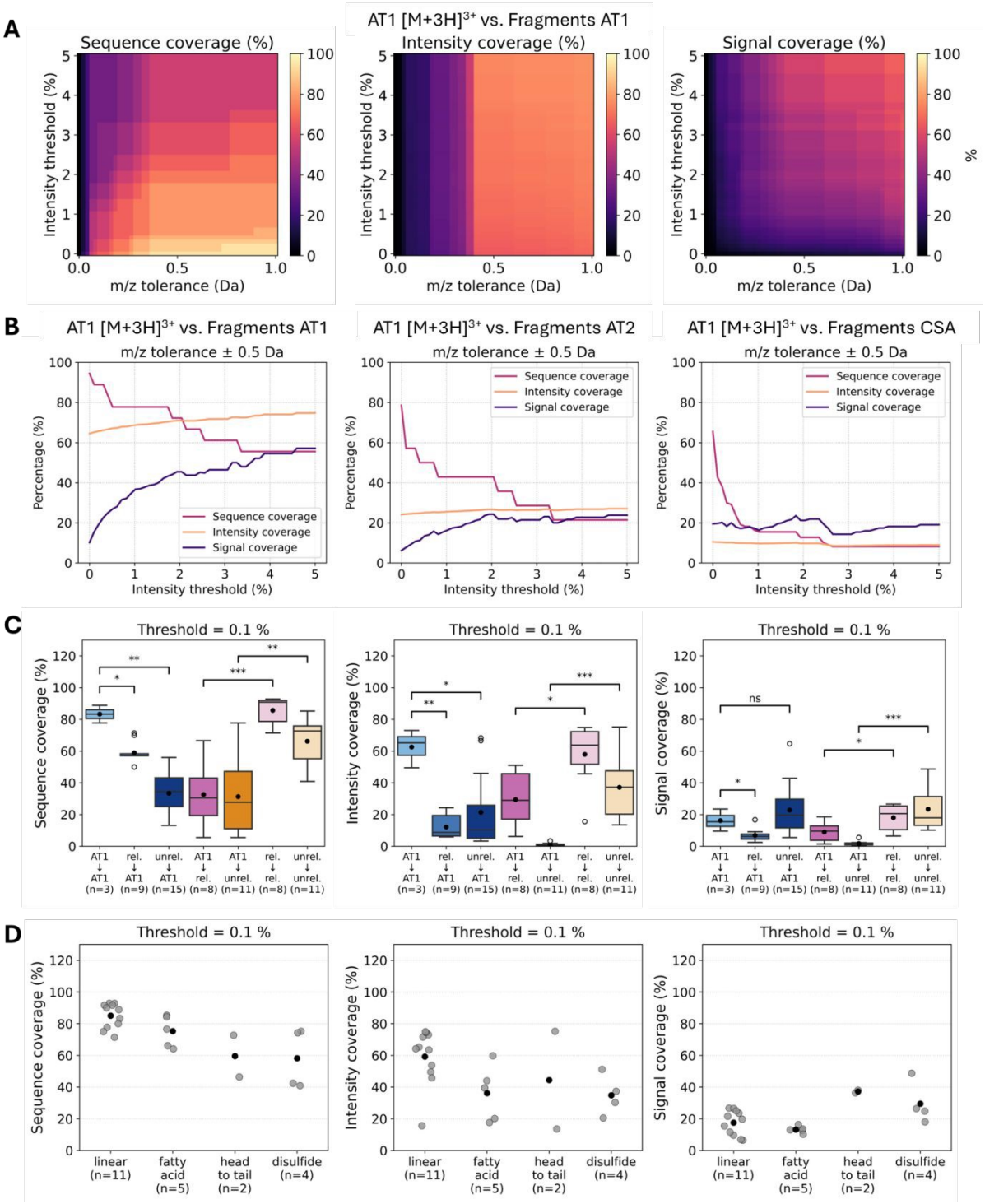
To investigate this further, we assessed whether the assignment metrics can reliably distinguish correct from incorrect matches across the full set of peptides including related and unrelated sequences. Using again AT1 as a reference peptide, we compared assignments in both directions: AT1 fragments assigned to spectra of related (AT2-4) or unrelated peptides (all other peptides in our set), and AT1 spectra assigned with fragments from related or unrelated peptides. Each charged state spectrum was treated as an independent data point. To evaluate the influence of the intensity threshold, we varied it across a range of values. At a representative threshold of ~~a~~ 0.1% (Figure 3C), average sequence coverage and intensity coverage were significantly higher for correct assignments than for related or unrelated matches. This trend was robust across a range of thresholds (0.01%, 0.5%, 1%, 5% shown in the Supporting information (S4)). Only-at the lowest threshold (0.01%) did sequence coverage fail to reach significance for the comparison of related fragments to AT1 spectra, which may reflect unspecific assignment of low-intensity ions. Signal coverage was in comparison to the other metrics less discriminative, particularly for related peptides and at low thresholds. Additionally, mean values of the signal coverage were generally lower than for the other metrics, indicating that many signals in the fragment spectra remain unexplained. Overall, these observations indicate that especially the sequence and intensity coverage metrics provide sufficient ability to distinguish correct from false assignments

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

in our dataset, even when comparing closely related peptides. It should be noted that these results were obtained using a relatively broad m/z tolerance of 0.5 Da. In addition, precursor mass was not used as a filtering criterion in this study, since low-resolution MS did not allow unambiguous molecular formula determination. With high-resolution instrumentation, narrower tolerances and precursor-based filtering could further improve the specificity of fragment assignments.



Analyst Accepted Manuscript

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

View Article Online
DOI: 10.1039/D6AN00162A

Figure 3: (A) Heatmap of sequence coverage, signal coverage, and intensity coverage as a function of *m/z* tolerance and intensity threshold for AT1 $[M+3H]^{3+}$. (B) Comparison of correct and incorrect assignments for AT1 $[M+3H]^{3+}$ at ± 0.5 *m/z*, using theoretical fragments of AT1 (correct), AT2 (related) and CSA (unrelated). (C) Assignment metrics for AT1 fragments assigned to AT1 spectra, related (rel.) spectra (AT2-4), and unrelated (unrel.) spectra, as well as assignments with calculated fragments of related and unrelated peptides to their respective spectra or to the AT1 spectra with 0.1% threshold and ± 0.5 *m/z*. Each charged state spectrum was considered an independent data point. Significance between correct and incorrect assignments was assessed by the Mann-Whitney U test (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). (D) Distribution of assignment metrics by peptide type at 0.1% threshold and ± 0.5 *m/z*. Individual data points are shown in grey and the mean value in black.

Finally, we assessed whether assignments for peptidomimetics were comparable to those of linear peptides. Figure 3D compares the different metrics between peptide types, with each charge state treated as an independent data point. On average, sequence and intensity coverage were lower for fatty acid-modified, cyclic, and disulfide-constrained peptides than for linear peptides. This difference was particularly pronounced for sequence coverage at higher thresholds as shown in the supporting information (S4). Intensity coverage remained more stable across thresholds, consistent with previous observations. Signal coverage showed no consistent trend. It might be used as a value reflecting the correlation between observed and assumed fragmentation pathways, but this requires further evaluation. Pairwise comparisons between related peptidomimetics, analogous to those described for AT1, further supported the discriminative power of the metrics. Correct and incorrect assignments were tested bidirectionally between LGL and SGL, between AT1 and CSA, and between SMT and OXY shown in the supporting information (S4). Although spectra and results varied across charge states, highlighting the importance of considering multiple metrics as well as charge states when available, correct assignments consistently produced higher sequence and intensity coverage across thresholds and charge states. When examining individual assignments, even correct assignment may yield in low absolute metric values (e.g. OXY $[M+H]^{+}$ and $[M+2H]^{2+}$). However, when the spectra were compared to related peptides, the correct assignments still resulted in higher values. This suggests that achieving an absolute threshold may not be necessary for correct identification and instead, aiming for a relative maximum in each case may be more appropriate. Nevertheless, to draw more conclusive results, larger datasets will need to be evaluated.

Case-specific evaluation of peptidomimetic fragmentation

Having established overall performance in peptide discrimination, we next applied the algorithm to specific peptidomimetic classes to explore their fragmentation behaviour in greater detail. First, we applied the algorithm to fatty acid-modified peptides to evaluate whether the inclusion of side-chain-specific fragments in assignment is reasonable and beneficial, as we found no prior reports addressing this. Since low charge states in the case of LGL and SGL exceeded the *m/z* 2000 detection limit and resulting fragment ions might as well, we mainly depended on highly charged fragments for our

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

evaluation. To account for reduced m/z resolution and increased spectral density at higher charge states, we applied a stricter tolerance of ±0.25 Da to reduce unspecific assignments. For comparison we focused on intensity coverage, as this metric was largely independent of the chosen threshold and is unaffected by changes in the number of calculated fragments when side-chain-specific fragments are included. Table 2 summarizes the intensity coverage for LGL and SGL across charge states, comparing assignments with and without fatty acid-specific fragments, along with the observed fragments and their intensities found. For LGL, inclusion of fatty acid-specific fragments increased intensity coverage for $[M+2H]^{2+}$ and $[M+3H]^{3+}$ by several percent, driven by fragments contributing >10% of the total intensity. A similar effect was observed for $[M+3H]^{3+}$ in SGL. In all cases, the M-fa2++ fragment was the most relevant fatty acid-specific fragment, corresponding to cleavage between the γ-Glu linker and the remaining peptide. These spectra for LGL and for SGL, together with the fatty acid side chain structure and fragmentation site, are shown in Figure 4. For the $[M+4H]^{4+}$ charge state, inclusion of fatty acid-specific fragments had little effect in LGL and none in SGL shown in the Supporting information (S3), underlining again the importance of considering multiple charge states. In general, assignments for the $[M+4H]^{4+}$ spectra yielded lower values, which may reflect limitations in precursor selection at high charges under low-resolution conditions but this would need to be investigated in more detail.

Open Access Article. Published on 30 March 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

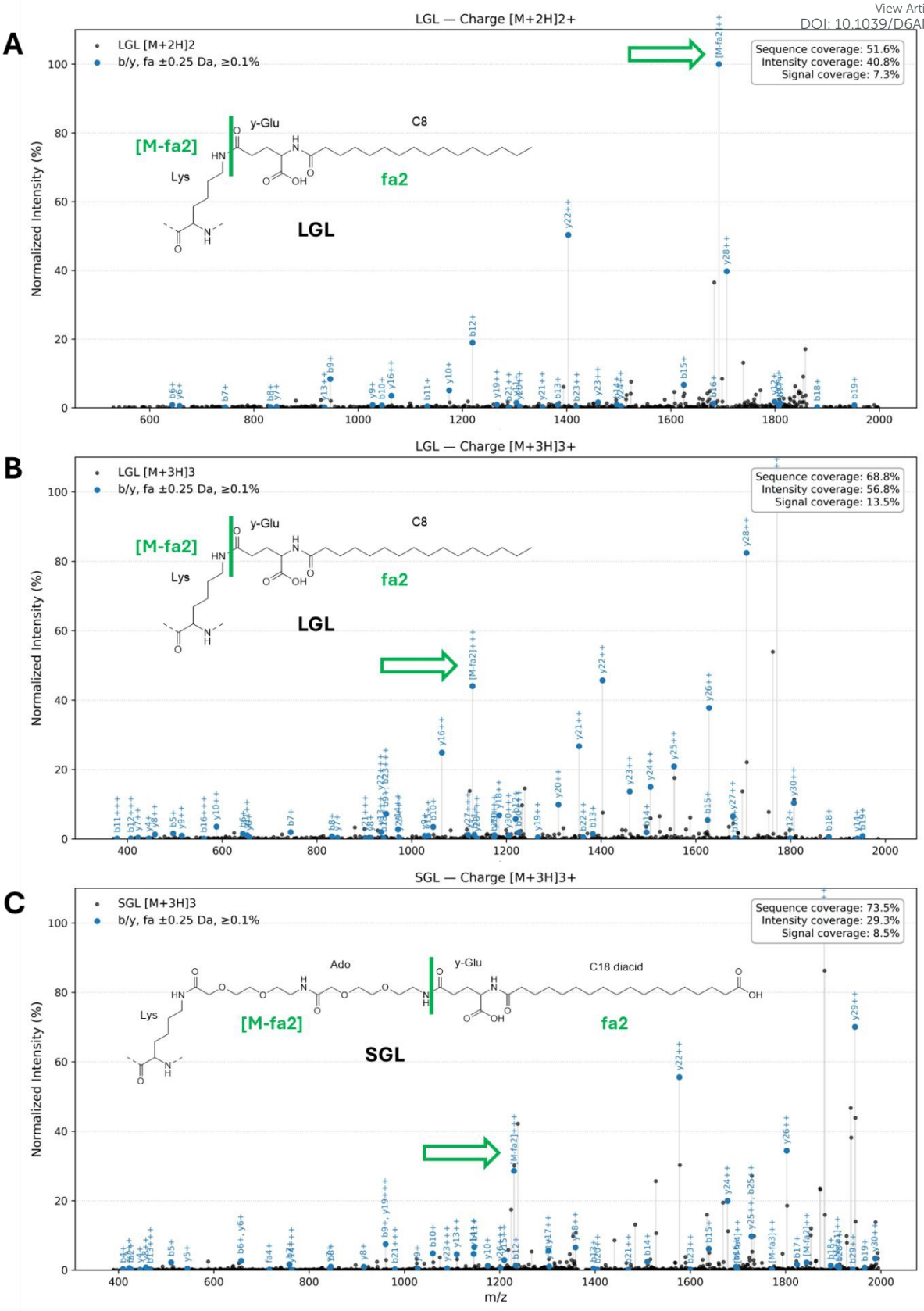


Figure 4: Assigned MS/MS spectra of (A) LGL $[M+2H]^{2+}$, (B) LGL $[M+3H]^{3+}$ and (C) SGL $[M+3H]^{3+}$. Side-chain fatty acid modifications are shown with the main fragmentation site [M-fa2] highlighted.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

Table 2: Intensity coverage for fatty acid-modified peptides SGL and LGL with and without inclusion of fatty acid-specific fragments (fa). Detected fragments are listed with their calculated mass, observed mass, and corresponding intensity found. Assignment was performed using a 0.1% intensity threshold and an *m/z* tolerance of ±0.25 Da.

Peptide	**Spectrum**	**Intensity coverage (%)**		**Fragments detected**	**Calculated mass (*m/z*)**	**Mass found (*m/z*)**	**Normalized Intensity (%)**
	Assignment	*Without fa fragments*	*with fa fragments*				
LGL	$[M+2H]^{2+}$	24.5	40.8	[M-fa2]++	1691.84	1692.07	100.0
	$[M+3H]^{3+}$	51.8	56.8	[M-fa2]+++	1128.23	1128.43	44.04
	$[M+4H]^{4+}$	7.9	8.0	Fa2 +	368.28	368.22	0.34
SGL	$[M+3H]^{3+}$	26.9	29.3	Fa2 +	426.28	426.50	0.14
				Fa4 +	716.43	716.68	0.14
				[M-fa1]++	1908.45	1908.41	1.10
				[M-fa2]++	1843.92	1844.15	2.20
				[M-fa2]+++	1229.62	1229.78	28.61
				[M-fa3]++	1771.39	1771.28	0.27
				[M-fa4]++	1698.85	1698.86	0.23
	$[M+4H]^{4+}$	11.3	11.3	None			

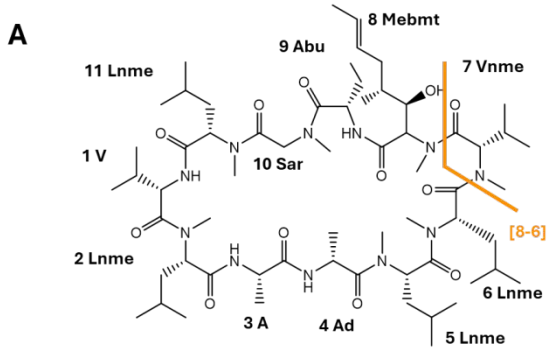
We next examined head-to-tail cyclized peptides in greater detail. To allow direct comparisons, MS/MS spectra for cyclic peptides were acquired under the same experimental parameters as for linear peptides, without specific adjustments. Since sequence-informative ions in cyclic peptides generally require two or more backbone cleavages, MS^n experiments are often considered advantageous for structural elucidation. To test this assumption, we compared assignment metrics for CSA across MS/MS (MS^2), MS^3, and MS^4 analysis. The structure of CSA with amino acid numbering is shown in Figure 5A. For this comparison, the $[M+2H]^{2+}$ precursor was used and the resulting sequence, intensity, and signal coverage are shown in Figure 5B. The obtained intensity coverage values were similar to those reported in the literature for cyclic peptides.[20] MS^3 improved sequence coverage by 8.7% compared to MS^2, but intensity coverage decreased by 6.9%, likely because additional fragmentation distributed signal intensity over a larger number of ions. MS^4 did not yield further improvement in sequence coverage over MS^3. Two-dimensional fragment maps of CSA (Figure 5C) illustrate the distribution of observed fragments and their relative intensities. The main fragment observed in MS^2 corresponded to [8-6], matching the initial loss of Vnme as previously described for $[CSA+Na]^+$ in MS^n analyses.[48] Notably, this visualization revealed differences in charge distribution: fragments detected with double charge were generally larger in size. This suggests that, in addition to intrinsic bond-cleavage probabilities, the ability of a fragment to retain charge influences whether it is ultimately observed in an MS/MS spectrum

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

and thus available for structural interpretation. Comparable trends were observed in MS^3 and MS^4 spectra, but no systematic changes indicating improved sequence-level interpretation at higher MS order were identified. Therefore, MS^n analysis may be more useful for the identification of individual fragments than for improving overall sequence coverage in cyclic peptides.



B

CSA $[M+2H]^{2+}$	MS^2	MS^3	MS^4
Sequence coverage (%)	72.7	80.9	80.9
Intensity coverage (%)	75.2	68.3	68.2
Signal coverage (%)	38.0	35.5	35.5

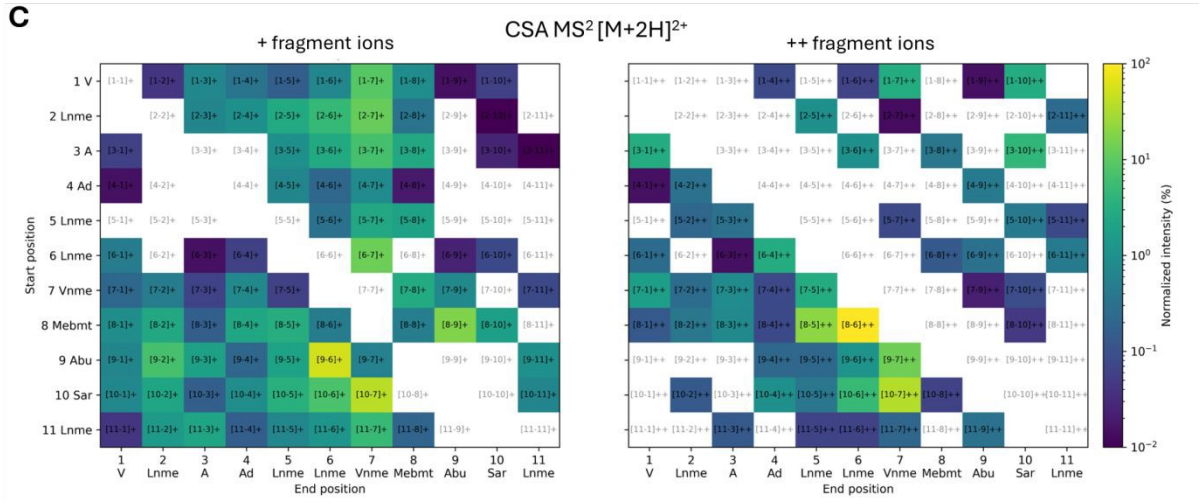


Figure 5: (A) Structures of CSA with amino acid numbering. (B) Sequence, intensity and signal coverage for MS^2, MS^3 and MS^4 of CSA $[M+2H]^{2+}$ (C) Two-dimensional fragment maps of CSA $[M+2H]^{2+}$ comparing singly and doubly charged ions found in MS^2 spectra and their normalized intensity.

Finally, we investigated disulfide-containing peptides in more detail. Disulfide bonds are known to cleave inefficiently under CID but can be more effectively addressed with ETD. To assess this, we compared spectra obtained from CID (CID-CID) alone to spectra resulting from ETD followed by CID (ETD-CID). Since ETD efficiency increases with precursor charge, the highest available charge state was used for each peptide.[49] For OXY, CID of $[M+2H]^{2+}$ yielded sequence, intensity, and signal coverages of 42.4%, 20.5%, and 24.8%, respectively. The ETD-CID spectrum showed lower coverage values (22.7%, 3.6%, and 12.5%). For SMT, CID of $[M+3H]^{3+}$ yielded sequence, intensity, and signal coverages of 74.3%, 51.2%, and 48.7%, respectively, while ETD-CID of $[M+2H]^{2+}$ showed values of

Analyst Accepted Manuscript

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

View Article Online
DOI: 10.1039/D6AN00162A

45.0%, 9.9%, and 26.2%. Thus, ETD-CID spectra displayed approximately half the sequence coverage with a fivefold lower intensity coverage compared to CID spectra, likely reflecting less efficient ETD fragmentation and the dominance of lower-charged precursor ions that remained unfragmented. According assigned spectra can be found in the Supporting information (S4). Figure 6 shows the structures of OXY and SMT with annotated bond types and the distribution of fragment types from CID and ETD-CID. In OXY, the CID spectrum was dominated by b/y-type fragments, with additional cyclic fragments and low-intensity disulfide-related ions. By contrast, ETD-CID produced a higher relative proportion of disulfide cleavage products, consistent with ETD-mediated S-S bond opening. In SMT, CID spectra were dominated by cyclic-related fragments with some disulfide products, whereas ETD-CID retained cyclic fragmentation as the main contribution but showed an increased ratio of disulfide-related fragments compared to CID. Together, these results demonstrate that the algorithm is useful not only for sequence assignment but also for investigating how different fragmentation methods influence peptidomimetic fragmentation pathways.

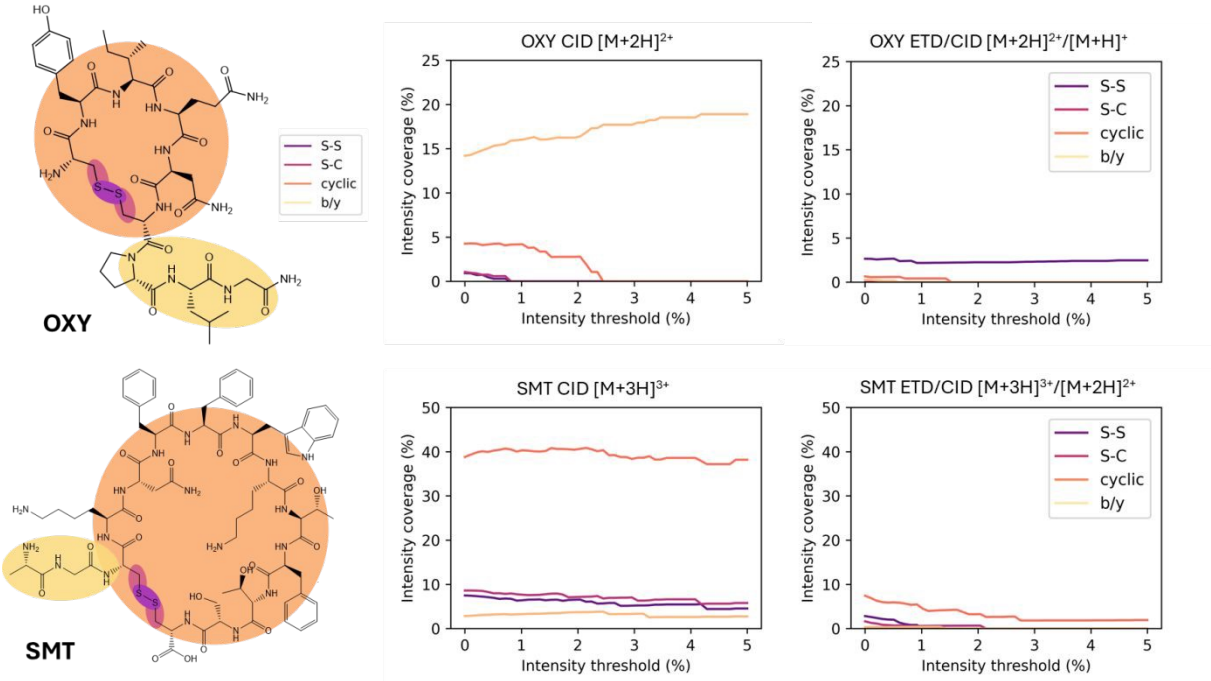


Figure 6: (Left) Structures of OXY and SMT with different types of bonds indicated. (Middle) CID fragmentation and (Right) ETD-CID fragmentation, showing intensity coverage by fragment type as a function of intensity threshold. Note the different y-axis ranges for OXY and SMT.

Conclusion and Outlook

In this work, we addressed tandem mass spectrometry fragmentation of peptidomimetics, taking structural features such as fatty acid conjugation, head-to-tail cyclization, and disulfide linkages into account. We demonstrated the utility of this structure-aware fragment calculation by integrating it with

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

the numerical assignment metrics sequence coverage, intensity coverage, and signal coverage for evaluation of MS/MS spectra. These metrics were sufficiently sensitive to discriminate correct from incorrect assignments, robust across related and unrelated peptides, and enabled systematic testing of assignment parameters in a structured and reproducible manner. Application to representative examples further highlighted both the potential and challenges of peptidomimetic fragmentation analysis. Fatty acid-modified peptides showed the benefit of including side-chain-specific fragments in assignment. Head-to-tail cyclic peptides illustrated how fragmentation preferences and charge retention shape sequence information, while disulfide-constrained peptides revealed shifts in fragmentation pathways from CID to ETD-CID conditions. These case studies demonstrate that the algorithm can support hypothesis-driven exploration of fragmentation behaviour in structurally diverse peptides within a reasonable timeframe for analysis. Although our dataset was limited, the framework provides a foundation for systematic evaluation of peptidomimetic MS/MS data based on numerical measures of assignment quality. Importantly, the approach is based on MS/MS spectral information and is broadly applicable across different analytical platforms and experimental setups. Future work may extend fragment calculations to additional modifications and more diverse peptide structures, include alternative fragmentation pathways, incorporate advanced assignment metrics, and explore variations in analytical parameters and fragmentation techniques to broaden applicability. By combining empirical experimentation with structure-aware fragment calculation, this approach contributes toward a more comprehensive and standardized MS/MS data analysis for complex peptidomimetics and biologically and pharmaceutically relevant peptides and paves the way for elucidation of previously unknown peptide-based structures.

Authors contributions

V.E. designed the study, programmed the algorithm, synthesized peptides, developed the MS/MS method, performed MS/MS analyses, evaluated the data, prepared the figures, and wrote the manuscript. A.M. performed MS/MS analyses and carried out proof-of-concept data analysis. A.S. synthesized peptides. C.S. provided resources and supervision. All authors read, reviewed and approved the final version of the manuscript.

Acknowledgements

The authors thank the laboratory of Prof. Dr. Gisbert Schneider for the access to the peptide synthesizer and the freeze drying facility.

References

(1) Henninot, A.; Collins, J. C.; Nuss, J. M. The Current State of Peptide Drug Discovery: Back to the Future? *J Med Chem* **2018**, *61* (4), 1382–1414. DOI: 10.1021/acs.jmedchem.7b00318.

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

(2) D'Aloisio, V.; Dognini, P.; Hutcheon, G. A.; Coxon, C. R. PepTherDia: database and structural composition analysis of approved peptide therapeutics and diagnostics. *Drug Discov Today* **2021**, *26* (6), 1409–1419. DOI: 10.1016/j.drudis.2021.02.019.
(3) Morrison, C. Constrained peptides' time to shine? *Nat Rev Drug Discov* **2018**, *17* (8), 531–533. DOI: 10.1038/nrd.2018.125.
(4) Lamers, C. Overcoming the Shortcomings of Peptide-Based Therapeutics. *Future Drug Discovery* **2022**, *4* (2), FDD75. DOI: 10.4155/fdd-2022-0005.
(5) Gare, C. L.; White, A. M.; Malins, L. R. From lead to market: chemical approaches to transform peptides into therapeutics. *Trends in Biochemical Sciences* **2025**, *50* (6), 467–480. DOI: https://doi.org/10.1016/j.tibs.2025.01.009.
(6) Erckes, V.; Steuer, C. A story of peptides, lipophilicity and chromatography – back and forth in time. *RSC Medicinal Chemistry* **2022**, *13* (6), 676–687, 10.1039/D2MD00027J. DOI: 10.1039/D2MD00027J.
(7) Nielsen, D. S.; Shepherd, N. E.; Xu, W.; Lucke, A. J.; Stoermer, M. J.; Fairlie, D. P. Orally Absorbed Cyclic Peptides. *Chem Rev* **2017**, *117* (12), 8094–8128. DOI: 10.1021/acs.chemrev.6b00838.
(8) Guo, T.; Steen, J. A.; Mann, M. Mass-spectrometry-based proteomics: from single cells to clinical applications. *Nature* **2025**, *638* (8052), 901–911. DOI: 10.1038/s41586-025-08584-0.
(9) Hellinger, R.; Sigurdsson, A.; Wu, W.; Romanova, E. V.; Li, L.; Sweedler, J. V.; Süssmuth, R. D.; Gruber, C. W. Peptidomics. *Nature Reviews Methods Primers* **2023**, *3* (1), 25. DOI: 10.1038/s43586-023-00205-2.
(10) Edman, P. Mechanism of the Phenyl Isothiocyanate Degradation of Peptides. *Nature* **1956**, *177* (4510), 667–668. DOI: 10.1038/177667b0.
(11) Elsayed, Y. Y.; Kühl, T.; Imhof, D. Edman Degradation Reveals Unequivocal Analysis of the Disulfide Connectivity in Peptides and Proteins. *Analytical Chemistry* **2024**, *96* (10), 4057–4066. DOI: 10.1021/acs.analchem.3c04229.
(12) Beck, J. G.; Frank, A. O.; Kessler, H. NMR of Peptides. In *NMR of Biomolecules*, 2012; 328–344.
(13) Schroeder, C. I.; Rosengren, K. J. Three-Dimensional Structure Determination of Peptides Using Solution Nuclear Magnetic Resonance Spectroscopy. In *Snake and Spider Toxins: Methods and Protocols*, Priel, A. Ed.; Springer US, 2020; 129–162.
(14) Spencer, R. K.; Nowick, J. S. A Newcomer's Guide to Peptide Crystallography. *Israel Journal of Chemistry* **2015**, *55* (6-7), 698–710. DOI: https://doi.org/10.1002/ijch.201400179.
(15) Steen, H.; Mann, M. The abc's (and xyz's) of peptide sequencing. *Nature Reviews Molecular Cell Biology* **2004**, *5* (9), 699–711. DOI: https://10.1038/nrm1468.
(16) Medzihradszky, K. F.; Chalkley, R. J. Lessons in de novo peptide sequencing by tandem mass spectrometry. *Mass Spectrom. Rev.* **2015**, *34* (1), 43–63. DOI: https://10.1002/mas.21406 From NLM Medline.
(17) Chi, H.; Chen, H.; He, K.; Wu, L.; Yang, B.; Sun, R. X.; Liu, J.; Zeng, W. F.; Song, C. Q.; He, S. M.; et al. pNovo+: de novo peptide sequencing using complementary HCD and ETD tandem mass spectra. *J Proteome Res* **2013**, *12* (2), 615–625. DOI: 10.1021/pr3006843 From NLM.
(18) Wiesner, J.; Premsler, T.; Sickmann, A. Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications. *Proteomics* **2008**, *8* (21), 4466–4483. DOI: 10.1002/pmic.200800329 From NLM.
(19) Townsend, C.; Furukawa, A.; Schwochert, J.; Pye, C. R.; Edmondson, Q.; Lokey, R. S. CycLS: Accurate, whole-library sequencing of cyclic peptides using tandem mass spectrometry. *Bioorg Med Chem* **2018**, *26* (6), 1232–1238. DOI: https://10.1016/j.bmc.2018.01.027 From NLM Medline.
(20) Liu, W.-T.; Ng, J.; Meluzzi, D.; Bandeira, N.; Gutierrez, M.; Simmons, T. L.; Schultz, A. W.; Linington, R. G.; Moore, B. S.; Gerwick, W. H.; et al. Interpretation of Tandem Mass Spectra Obtained from Cyclic Nonribosomal Peptides. *Anal. Chem.* **2009**, *81* (11), 4200–4209. DOI: https://10.1021/ac900114t.
(21) Bittremieux, W.; Ananth, V.; Fondrie, W. E.; Melendez, C.; Pominova, M.; Sanders, J.; Wen, B.; Yilmaz, M.; Noble, W. S. Deep Learning Methods for De Novo Peptide Sequencing. *Mass*

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

Spectrometry Reviews **2024**, *n/a* (n/a). DOI: https://doi.org/10.1002/mas.21919 (accessed 2025/08/25).
(22) Eng, J. K.; McCormack, A. L.; Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* **1994**, *5* (11), 976–989. DOI: 10.1016/1044-0305(94)80016-2 From NLM.
(23) Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20* (18), 3551–3567. DOI: 10.1002/(sici)1522-2683(19991201)20:18<3551::Aid-elps3551>3.0.Co;2-2 From NLM.
(24) Craig, R.; Beavis, R. C. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **2004**, *20* (9), 1466–1467. DOI: 10.1093/bioinformatics/bth092 From NLM.
(25) Tyanova, S.; Temu, T.; Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **2016**, *11* (12), 2301–2319. DOI: 10.1038/nprot.2016.136.
(26) Kim, S.; Pevzner, P. A. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nature Communications* **2014**, *5* (1), 5277. DOI: 10.1038/ncomms6277.
(27) Yilmaz, Ş.; Vandermarliere, E.; Martens, L. Methods to Calculate Spectrum Similarity. In *Proteome Bioinformatics*, Keerthikumar, S., Mathivanan, S. Eds.; Springer New York, 2017; 75–100.
(28) Ma, B.; Zhang, K.; Hendrie, C.; Liang, C.; Li, M.; Doherty-Kirby, A.; Lajoie, G. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun Mass Spectrom* **2003**, *17* (20), 2337–2342. DOI: 10.1002/rcm.1196 From NLM.
(29) Ma, B. Novor: Real-Time Peptide de Novo Sequencing Software. *Journal of the American Society for Mass Spectrometry* **2015**, *26* (11), 1885–1894. DOI: 10.1007/s13361-015-1204-0.
(30) Tran, N. H.; Zhang, X.; Xin, L.; Shan, B.; Li, M. De novo peptide sequencing by deep learning. *Proceedings of the National Academy of Sciences* **2017**, *114* (31), 8247–8252. DOI: 10.1073/pnas.1705691114 (accessed 2025/08/25).
(31) Eloff, K.; Kalogeropoulos, K.; Mabona, A.; Morell, O.; Catzel, R.; Rivera-de-Torre, E.; Berg Jespersen, J.; Williams, W.; van Beljouw, S. P. B.; Skwark, M. J.; et al. InstaNovo enables diffusion-powered de novo peptide sequencing in large-scale proteomics experiments. *Nature Machine Intelligence* **2025**, *7* (4), 565–579. DOI: 10.1038/s42256-025-01019-5.
(32) Hayashi, A.; Goto, Y.; Saito, Y.; Suga, H.; Morimoto, J.; Sando, S. Oxidation-guided and collision-induced linearization assists de novo sequencing of thioether macrocyclic peptides. *Chemical Communications* **2024**, *60* (70), 9436–9439, 10.1039/D4CC03179B. DOI: 10.1039/D4CC03179B.
(33) Cole, S. R.; Ma, X.; Zhang, X.; Xia, Y. Electron transfer dissociation (ETD) of peptides containing intrachain disulfide bonds. *J Am Soc Mass Spectrom* **2012**, *23* (2), 310–320. DOI: https://10.1007/s13361-011-0300-z From NLM Medline.
(34) Li, H.; O'Connor, P. B. Electron capture dissociation of disulfide, sulfur-selenium, and diselenide bound peptides. *J Am Soc Mass Spectrom* **2012**, *23* (11), 2001–2010. DOI: https://10.1007/s13361-012-0473-0 From NLM Medline.
(35) Xu, H.; Zhang, L.; Freitas, M. A. Identification and Characterization of Disulfide Bonds in Proteins and Peptides from Tandem MS Data by Use of the MassMatrix MS/MS Search Engine. *Journal of Proteome Research* **2008**, *7* (1), 138–144. DOI: 10.1021/pr070363z.
(36) Bern, M.; Kil, Y. J.; Becker, C. Byonic: Advanced Peptide and Protein Identification Software. *Current Protocols in Bioinformatics* **2012**, *40* (1), 13.20.11–13.20.14. DOI: https://doi.org/10.1002/0471250953.bi1320s40.
(37) Liu, M.-Q.; Zeng, W.-F.; Fang, P.; Cao, W.-Q.; Liu, C.; Yan, G.-Q.; Zhang, Y.; Peng, C.; Wu, J.-Q.; Zhang, X.-J.; et al. pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification. *Nature Communications* **2017**, *8* (1), 438. DOI: 10.1038/s41467-017-00535-2.

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

(38) Erckes, V.; Hilleke, M.; Isert, C.; Steuer, C. PICKAPEP: An application for parameter calculation and visualization of cyclized and modified peptidomimetics. *Journal of Peptide Science* **2024**, *30* (12), e3646. DOI: https://doi.org/10.1002/psc.3646.
(39) *The Python Language Reference*. https://docs.python.org/3/reference/ (accessed January 22, 2024).
(40) Landrum, G. *RDKit: Open-Source Cheminformatics*. 2023. https://www.rdkit.org/ (accessed January 22, 2024).
(41) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; et al. Array programming with NumPy. *Nature* **2020**, *585* (7825), 357–362. DOI: 10.1038/s41586-020-2649-2.
(42) *pandas User Guide*. https://pandas.pydata.org/ (accessed January 22, 2024).
(43) *SciPy*. https://scipy.org/ (accessed 23.08.2024).
(44) *Matplotlib*. https://matplotlib.org/ (accessed 23.08.24).
(45) *Seaborn*. https://seaborn.pydata.org/ (accessed 12.07.2025).
(46) Erckes, V.; Hilleke, M.; Isert, C.; Steuer, C. PICKAPEP: An application for parameter calculation and visualization of cyclized and modified peptidomimetics. *J. Pept. Sci.* **2024**, e3646. DOI: https://10.1002/psc.3646 From NLM Publisher.
(47) Heissel, S.; He, Y.; Jankevics, A.; Shi, Y.; Molina, H.; Viner, R.; Scheltema, R. A. Fast and Accurate Disulfide Bridge Detection. *Molecular & Cellular Proteomics* **2024**, *23* (5), 100759. DOI: https://doi.org/10.1016/j.mcpro.2024.100759.
(48) Ahn, E. Y.; Shrestha, A.; Hoang, N. H.; Huong, N. L.; Yoon, Y. J.; Park, J. W. Structural characterization of cyclosporin A, C and microbial bio-transformed cyclosporin A analog AM6 using HPLC-ESI-ion trap-mass spectrometry. *Talanta* **2014**, *123*, 89–94. DOI: 10.1016/j.talanta.2014.01.067.
(49) Hellinger, J.; Brodbelt, J. S. Impact of Charge State on Characterization of Large Middle-Down Sized Peptides by Tandem Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2024**, *35* (8), 1647–1656. DOI: 10.1021/jasms.3c00405.

Open Access Article. Published on 20 May 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript

View Article Online
DOI: 10.1039/D6AN00162A

Data availability statement:

The data supporting this article have been included as part of the supplementary information (SI). Supplementary information is available. Other data are available from the corresponding authors upon request.

Open Access Article. Published on 30 March 2026. Downloaded on 5/21/2026 9:47:19 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Analyst Accepted Manuscript