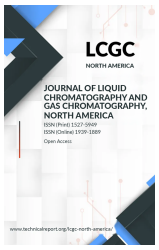


Journal of Liquid Chromatography and Gas Chromatography, North America
https://technicalreport.org/lcgcnam

LC GC N. Am., Vol. 43, No.3, 28-34, 2025

Technical Reports Press

RESEARCH ARTICLE

Interference-Staged Charge-State Evidence Partitioning for MALDI MS/MS Identification of Peptides, Intact Proteins, and Lysate Components

Louis E. Brus[1*]

[2]*Department of Chemistry, Columbia University, New York, New York 10027, USA*
[*]*Corresponding author*

Corresponding author:
leb26@columbia.edu

Abstract
Co-isolation in MALDI tandem mass spectrometry is conventionally considered spectral contamination; yet, the acquisition information can contain evidence of charge-state continuity, shared precursor windows, and analyte-specific fragment retention. In this paper, we introduce Charge-State Evidence Partitioning (CSEP) as an interference-aware analytical workflow to discriminate between an ambiguous overlap that requires spectrum rejection as an unresolved case and a meaningful multi-analyte measurement with recoverable fragments. The present investigation uses seven channels for analytes that span three levels of complexity: a pair of ACTH 1-17 and angiotensin II peptides, a pair of ubiquitin and cytochrome C proteins, and a tripartite mixture of bradykinin, myoglobin, and *Escherichia coli* lysate. The research question is formulated narrowly: can evidence from a shared precursor window be routed into the relevant charge-state family and used to search the database only if its location suggests such an operation to be justified? CSEP applies a window registry, charge-state routing, fragment recurrence evaluation, and cross-channel evidence specificity test before database search results are generated. The score outcome depended on the analyte: ACTH 1-17 from 71 up to 153, angiotensin II from 25 up to 66, ubiquitin from 42 up to 288, bradykinin from 3 up to 56, and DNA-binding protein HU-alpha from 55 up to 254; the cytochrome C and myoglobin were considered separate protein channels with respective scores of 225 and 149. Analyte score improvement was significant especially in cases when the protection against contamination provided by several charge-state windows enabled detection of a weak fragment signal in a crowded mixed window.

Keywords: MALDI MS/MS; chimeric spectra; charge-state continuity; precursor co-isolation; top-down proteomics; peptide identification; protein identification; analytical spectroscopy.

Received: 12 April 2025
Accepted: 09 August 2025
Available online: 30 September 2025

1. Introduction

In MS/MS tandem mass spectrometry experiment, a connection between an event generated by a particular instrument and molecular outcomes is considered. One ion is chosen as the parent ion, fragmented, and product ions are correlated to a sequence of peptides or proteins. In ideal case, the precursor window consists of one single analyte; however, in reality there might be more than one. Practically, a tandem spectrum is chimeric, containing product ions of both target analyte and others co-isolated in the mass window, as well as background signals, matrix effects, neutral loss chemistry, internal fragment formation, and so forth. In turn, from the analytical point of view, this means that besides the problem of assigning product ions to particular peaks, there is an issue of determining which peaks correspond to an analyte and which do not. A number of detected peaks would make a search algorithm reach valid matches, as well as low-score results or even false positives.

Historically, MALDI technique was highly appreciated for providing relatively clean ions and facilitating the direct study of proteins. It was developed simultaneously with electrospray ionization, giving two complementary sources of ions for biomolecule mass spectrometry: desorption with pulsed laser and ionization from the solution forming multiply charged ions [1, 2]. Progress in atmospheric pressure and liquid-supported MALDI techniques made the separation harder, since

Cite as: L. E. Brus (2025). Interference-Staged Charge-State Evidence Partitioning for MALDI MS/MS Identification of Peptides, Intact Proteins, and Lysate Components. LC GC N. Am., 43(3) (2025) 28-34.

it was shown that multiply charged peptides and proteins could be generated in the latter case [3,4]. The benefits of multiply charged ions lie in increasing the m/z range of analytes available for investigation and the ability of large-sized molecules to fragment effectively. At the same time, multiple charging introduces complexity to precursor ion landscape, since a peptide or protein will exist as a set of charge state-related ions, rather than a single precursor.

As far as the problem of co-isolated ions in peptides and proteins identification is concerned, it is important to remember that the assumption behind the use of a database search procedure is a certain level of purity of submitted ions. Thus, if ions in some of the product ions originate from another molecule, valuable data will be diminished by the interference. Research shows that co-isolated MS/MS spectra decrease peptide identification efficiency and distort results of such procedures [5,6]. The possibility of using computational approaches to overcome the issue has been introduced: mixed-spectrum searching and probability theory-based methods utilize multi-analyte MS/MS information, extracting it appropriately [7, 8]. The proposed work follows this idea; however, it uses charge correlation criterion instead of chromatography.

There are several currently known ways to solve the co-isolation problem: iterative and multipeptide approaches and isolation window repetition, library use, and pseudo-MS/MS reconstruction help to determine the origins of precursors [9, 10]. Recently, applications of data-independent acquisition techniques have been reviewed and highlighted due to their ability to preserve fragment information between isolation windows [11, 12]. In addition, the problem of protein identification in top-down MS/MS experiments suffers from complicated charge envelopes, isotopic distributions, and many possible fragment types. There exist programs such as TopPIC; however, the problem of interference remains a limiting factor in mixing spectra [13, 14].

All in all, a rich background on co-isolated spectra, related problems, and potential solutions can be seen in the current literature. On the contrary, the charge state continuity phenomenon in MALDI MS/MS spectra has not received enough attention in spite of having great potential. Instead of a charge state being an artifact of liquid- and atmospheric pressure-based techniques, it becomes a powerful tool in the form of a charge state family for determining if two precursor windows belong to the same or different precursors. Consequently, a notion of the interference stage in Charge State Evidence Partitioning arises.

Research questions posed here are relatively narrow and specifically focused on the problem discussed: does the location of shared precursor windows in the charge state family transform ambiguous or low evidence ion-product into reliable peptides/proteins identification evidence? The main idea differs from the statement that the application of evidence partitioning helps to obtain higher scores in that the location of improvement is specified as well as in which types of ions the effect can be observed. Also, whether or not the effect will be changing during transition from a two-peptide mixture to intact proteins, and then to a mixture of peptides, proteins, and lysate will be assessed.

To address the research question, this paper proposes the interference stage MALDI MS/MS dataset. Its lower stratum consists of two peptides with shared precursor windows at m/z 524 and 1047. The middle stratum is represented by two intact proteins with overlapping charge state families at m/z 952. Finally, the upper stratum consists of bradykinin, myoglobin, and bacterial lysate signal, overlapping at m/z 1060. Such stratification serves both as an increasing step-wise complexity of chemical compounds involved, as well as an effective way of demonstrating the power of evidence partitioning based on charge state families to separate interfering ions.

2. Materials and methodology

2.1. Interference-staged MALDI MS/MS dataset

Variables for analysis involve the type of analyte, location of the precursor window, assignment of charge states, sample preparation parameters, instrumental conditions, and score results of database search in regard to precursor ion connectivity [15]. Variables are stratified so that peptide-peptide, intact-protein, and peptide-protein-lysate interactions could be analyzed consistently based on evidence routing.

Peptides stratum: ACTH 1–17 and angiotensin II. Intact proteins stratum: bovine ubiquitin and equine cytochrome C. Complex stratum: bradykinin, equine myoglobin, and a bacterial lysate component, named DNA-binding protein HU-alpha. Bacterial lysate components were precipitated with trichloroacetic acid, washed with acetone, and purified with C18 ZipTip, eluting in acetonitrile/water. The concentrations of standard analytes were 10 micromolar, and mixed in the way that the same intact-protein signal levels were achieved. Aqueous liquid supported MALDI matrix was alpha-cyano-4-hydroxycinnamic acid, water/acetonitrile, and ethylene glycol. Positive ion measurements were carried out at m/z 100-2000 and collision induced dissociations at 30-60 volts.

Table 1 defines the measurement record as described in the stratum summary, based on types of interferences rather than names of samples alone, which provides the basis for the experiment design. Stratum one presents the smallest peptide-overlap system, stratum two presents a test case at higher masses with wider charge-state envelope systems, and stratum three provides a crowded environment where peptides and proteins will be competing against each other for detection. This allows for both explanation of score increases as well as understanding the different score responses by analyte types.

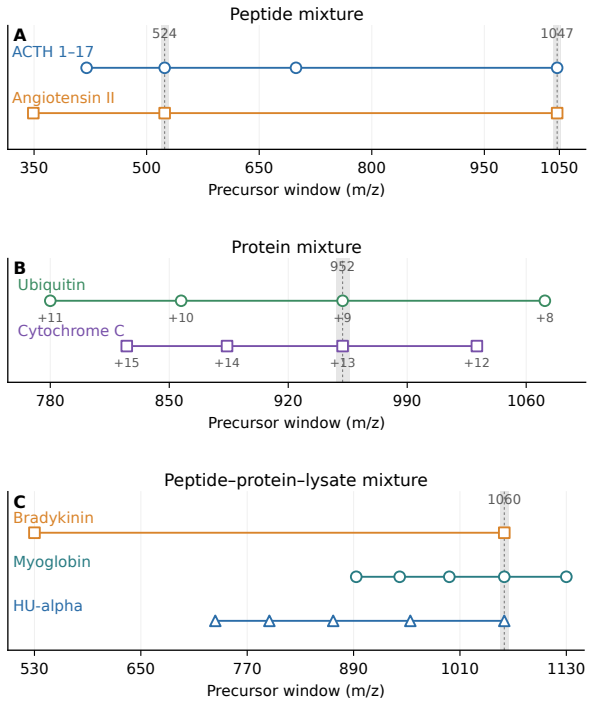


Figure 1: Precursor-window map for the three interference strata.

The precursor topology illustrated in Fig. 1 is an instance of the physical measure problem prior to the interpretation of score results. Given that the peptide consists of two separate shared regions, the issue at stake is how to maintain the information about angiotensin II without burdening the ACTH 1-17 channel with too much data. In other words,

Table 1: Interference-staged MALDI MS/MS dataset.

Stratum	Analyte channels	Shared or diagnostic windows	Role in the study
Peptide overlap	ACTH 1–17; angiotensin II	m/z 349, 420, 524, 699, and 1047; shared regions at 524 and 1047	Tests whether two peptide channels can both retain useful evidence when the same precursor region contributes to more than one assignment.
Protein-standard overlap	Bovine ubiquitin; equine cytochrome C	Ubiquitin +8 to +11; cytochrome C +12 to +15; shared region near m/z 952	Tests whether broad intact-protein charge envelopes can be separated by charge-family routing rather than by choosing only one apparently clean window.
Peptide–protein–lysate overlap	Bradykinin; equine myoglobin; *E. coli* HU-alpha	Bradykinin m/z 530 and 1060; myoglobin m/z 893, 942, 998, 1060, and 1130; HU-alpha m/z 734, 795, 867, 954, and 1060	Tests whether a small peptide, a spiked protein, and a bacterial protein can be interpreted when all three pass through the m/z 1060 region.

there is a need for charge series interpretation instead of just isolating precursor windows. The complex problem is definitely the most difficult to solve as m/z 1060 corresponds to the intersection of bradykinin, myoglobin, and HU-alpha from bacteria.

2.2. CSEP evidence-routing procedure

In the case of this current study, CSEP takes the form of an evidence routing mechanism, rather than being a black box for scoring. First, the selected windows must be registered in terms of the precursor-channel relationship with respect to the putative analyte channels. Next comes the charge-series routing, through which each window will be interpreted in the light of the corresponding charge-state sequence, instead of just the separate spectra. Recurrence review is the third step whereby the recurrent product ions from the charge family routes are saved as good sequencing evidence. Specificity review comes next to downgrade competing ions prior to the database search process.

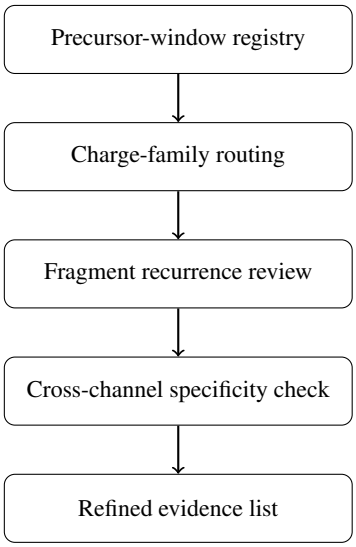


Figure 2: Evidence-routing sequence used in CSEP.

CSEP uses a procedure for determining which routes to use when analyzing MS/MS spectra shown in Figure 2. While peak list concatenation uses the concatenated peak lists of retained fragments as evidence that the fragmentation is occurring correctly, CSEP checks whether there is any evidence of the same in the appropriate charge family of the fragments. Moreover, it is necessary that the specificity of the fragments is taken into account, as a fragment appearing repeatedly without a channel assignment is not as helpful as a fragment repeatedly appearing in one specific analyte.

2.3. Analytical reporting strategy

This format of reporting has been specifically designed with respect to the research question posed by the paper. The relevant numeric information has been obtained from the score obtained from either a mixed window or a single window, and that from a score from a charge state evidenced by CSEP. The interpretation of data will be conducted keeping in mind if the score was obtained from a channel based on a peptide with only two or three useful windows, or charge states of proteins, or channels with several contributions. Figures will be used to convey the context of the data presented, and tables to represent numeric data.

3. Results and discussion

3.1. Overall evidence structure

It can be observed from the trend in scores that, in situations where the mixed window is not isolated, the importance of the continuity of charge states increases. The improvement obtained for ACTH 1-17, angiotensin II, ubiquitin, bradykinin, and HU-alpha molecules has been relative to the comparisons made to determine the percentage improvement, although the amount of improvement in different molecules varied greatly. The greatest percentage improvement obtained was for bradykinin due to extremely poor scores for its mixed window. However, the maximum absolute improvements have been obtained for ubiquitin and HU-alpha molecules due to adequate repeating information in multi-windows.

The evidence ledger of Figure 3 makes the connection clear between the precursor shape and the number of peaks generated. The membership matrix indicates precisely which windows must be interpreted rather than dismissed. It becomes apparent from the grouping ledger that while it may be possible for a single m/z space to house multiple peaks, this can be done only when there is consistency among their corresponding charge family evidence. From the audit and scores panels, it can be seen clearly that the confidence of identification of the peptide cannot be based on the presence of a single common window but should depend on how it behaves in the entire process of evidence collection.

The results presented in Table 2 provide an answer to the numerical portion of the research question. It is clear that CSEP did not generate a standard, cosmetic increase, but rather altered the interpretation of various windows. The angiotensin-II window shifted from having a nonsignificant mixed window score to a significant peptide score. The ubiquitin window shifted from having a nonsignificant mixed window score to being the window with the highest final score. Bradykinin was still lower numerically compared to the intact-protein windows, but the shift from 3 to 56 made all the difference, since it proved the

Table 2: Identification scores after evidence partitioning.

Channel	Assigned identity	CSEP evidence windows	Comparator	CSEP	Gain
ACTH 1–17	P01189 sequence region	420, 524, 699, 1047	71 mixed; 122 single	153	+82 vs mixed
Angiotensin II	P01019 sequence region	349, 524, 1047	25* mixed	66	+41
Ubiquitin	Bovine ubiquitin	+8 to +11 charge-state region	42* mixed; 146 single	288	+246 vs mixed
Cytochrome C	Equine cytochrome C, P00004	+12 to +15 charge-state region	–	225	–
Bradykinin	Human kininogen region, P01042	530, 1060	3* mixed	56	+53
Myoglobin	Equine myoglobin, P68082	893, 942, 998, 1060, 1130	–	149	–
HU-alpha	*E. coli* HU-alpha, P0ACF0	734, 795, 867, 954, 1060	55 single	254	+199 vs single

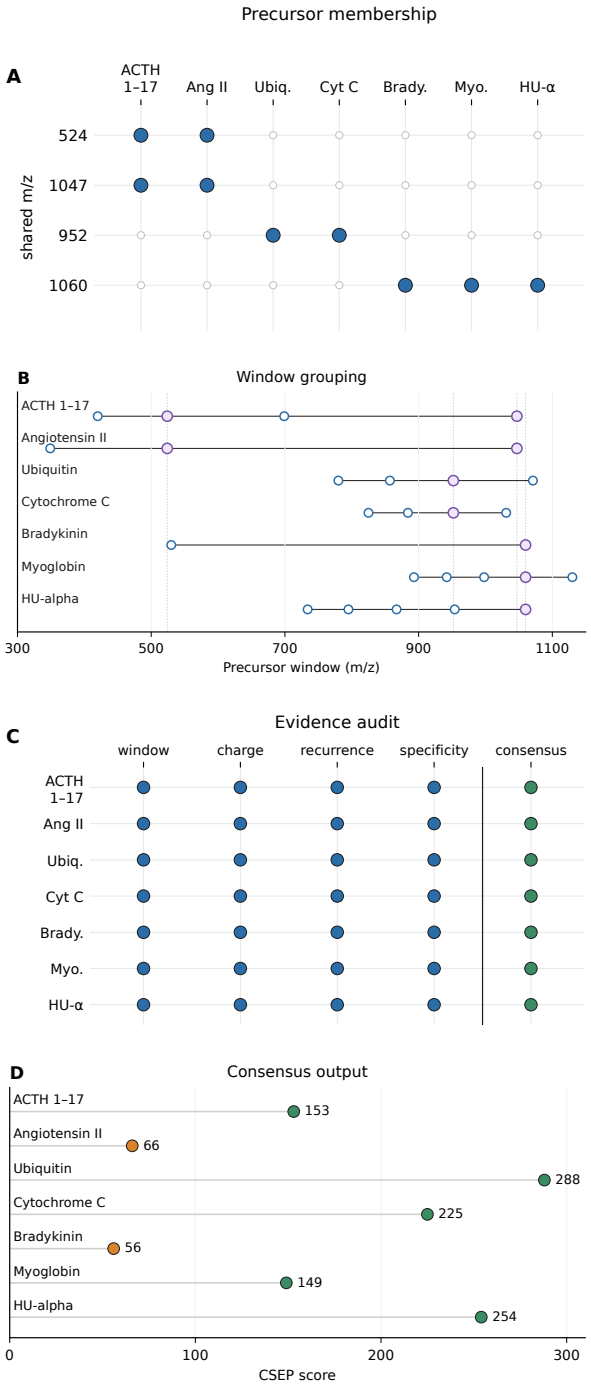


Figure 3: Evidence ledger linking shared windows to final analyte channels.

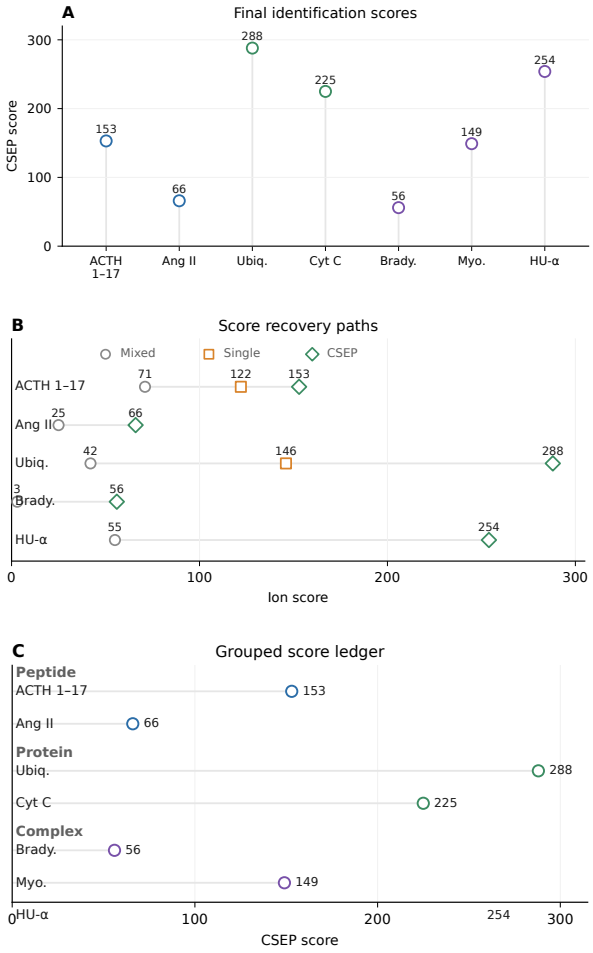


Figure 4: Final score response across the seven analyte channels.

analytical point that even a weak peptide can remain in its crowded precursors if the charge correlation between two windows is particular. Peptide rescue from Fig.4 refers to the output classification based on response type and is demonstrated by angiotensin II and bradykinin samples where the score result for both peptides is average, yet the transition from mixed response to peptide rescue response is critical enough to change the score evaluation. Protein reinforcement is the second output response type represented by the case of ubiquitin and HU-alpha proteins in which many routes resulted in a significant score for each of them. Finally, channel resolution represents a unique approach to interpreting output based on a channel creation for an analyte in mixed environment without evaluating two scores for the peptide. This is demonstrated by the cases of cytochrome C and myoglobin peptides in this study.

3.2. Peptide overlap: ACTH 1–17 and angiotensin II

Peptide layer may seem to provide the simplest test, but this test cannot be called the least significant one. There are three reasons why common areas could appear on m/z of 524 and 1047: contamination, data integration as valuable information, or handling common area for

the analyte under examination. The last one gives the best explanation for score results for these two peptides: four-window channel was opened for ACTH 1-17, and angiotensin II was found in a three-window channel.

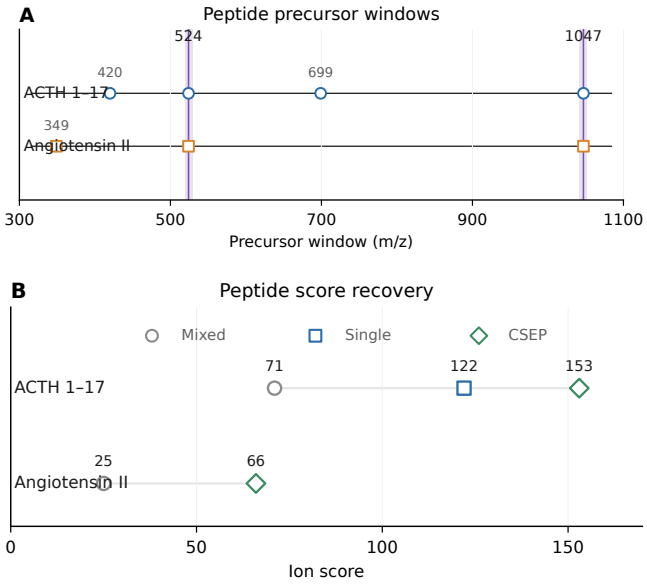


Figure 5: Peptide-window routing and score recovery.

Peptide partition in Fig. 5 indicates that the use of the shared peptides windows will benefit only in particular cases. For example, considering ACTH 1–17, it is clear that using the windows separately gives better results as the total score is higher (153 against 71 for the mixed windows and 122 for a single window). So, here the windows add some value to the peptide recognition. On the other hand, for angiotensin II, although the total score of 25 is not enough to report the finding reliably, the score after routing (66) proves that the sequence specificity is preserved even in the small window. Thus, the answer to one of our questions is: the windows at m/z 524 and 1047 are recoverable because each of them is used in the context of charge states.

3.3. Protein-standard overlap: ubiquitin and cytochrome C

The Protein Standard Layer looks into the applicability of the approach when the distribution of the studied compounds becomes more diverse. The problem of intact protein makes tandem identification hard due to fragmentation density, the spread of the charge states across different precursors, and the possibility of incomplete coverage of an amino acid chain. CSEP may come in handy in this situation, since no clarity for all the windows is required for it.

As per Figure 6, discrimination between ubiquitin and cytochrome C does not happen due to similar m/z choices, since the charge families have become closer at the mixed region. Essentially, it turns into a task of sorting out what of the fragment-ion information belongs to the charge family. For instance, the score value of ubiquitin went from being relatively insignificant in the mixed window to 288 by traversing through the charge families from +8 to +11. What is more important, the final score value was higher than the one in single window comparison of 146. It shows that charge family consensus could beat a clean spectrum. As for cytochrome C, it acquired a score value of 225 by traversing through the +12 to +15 region. It is quite clear that discrimination has happened since two different channels of intact proteins have been successfully separated in a single mass measurement stratum.

With regard to the top-down and middle-down approach, the achieved result can be described as a significant one within the framework of

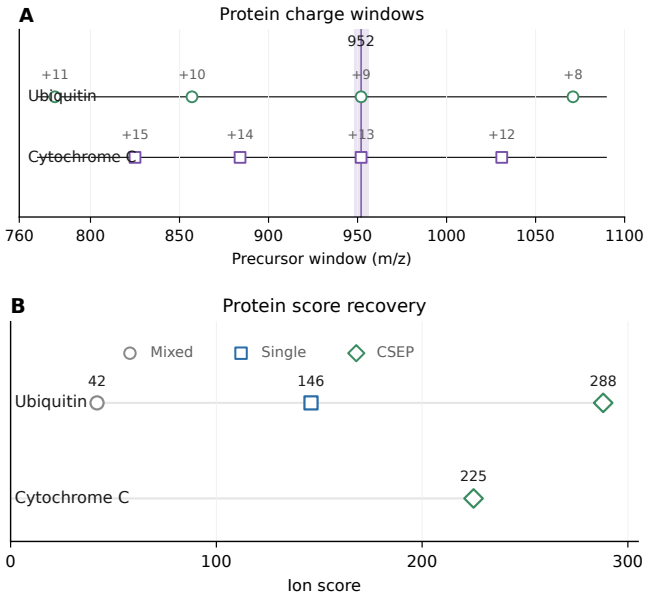


Figure 6: Protein charge-family partitioning.

partial mixing. In fact, one may try to be conservative enough to get rid of such windows to make sure that no misleading data has been generated. This strategy is safe but quite incomplete. Specifically, when the windows provide little sequence of intact proteins, the windows should remain conditional.

3.4. Peptide–protein–lysate overlap around m/z 1060

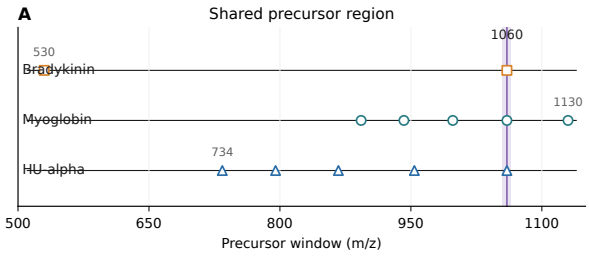
Complex Stratum This can be said about the region of m/z 1060. In fact, it could easily have been considered crowded evidence in case it stood alone. However, in CSEP it is regarded as controversial evidence that is dependent on the adjacent windows.

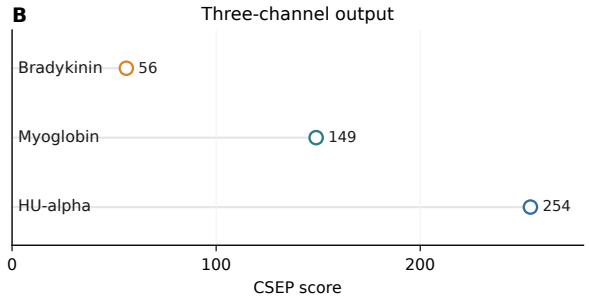
This explains why the case with lysate is not a mere noisy analogy of either peptides or proteins. Bradykinin has only two routed windows, at m/z 530 and 1060, which implies that the recovery of this substance is contingent upon the specificity of evidence obtained in the shared region. At the same time, both myoglobin and HU-alpha contain multiple windows, thus offering the possibility of assessing the substance in the light of charge-state evidence pattern. The resulting improvement in the scores is impressive: bradykinin goes from 3 to 56; myoglobin, up to 149; and HU-alpha, from 55 to 254. This illustrates how CSEP can save an analyte with low evidence and resolve other neighboring proteins.

Recovery of HU-alpha is significant because it indicates the possibility to use non-purified samples. A bacterial protein channel was recovered despite the background noise and the presence of fragmented ions in a single lysate measurement. The m/z 1060 region can contribute data only in combination with other windows, m/z 734, 795, 867, and 954 of the same substance. In simple terms, the shared window was not the only reason for the recovery of HU-alpha.

3.5. Comparison with established interpretation strategies

The results position CSEP between two well-known approaches. An approach that strictly rejects overlapping windows helps avoid false discoveries but discards valuable information by relying merely on the superiority of routed multi-windows over one-windowed interpretation of ACTH 1–17 and ubiquitin. In turn, an approach that allows a free mixture of data offers to submit more peaklists for database searching. However, this approach also creates risks of cross-channel contaminations. CSEP is special because it admits a window into a mixture only





C

Window and score summary

Analyte	Windows	Score
Bradykinin	530, 1060	56
Myoglobin	893–1130	149
HU-alpha	734–1060	254

Figure 7: Three-channel evidence recovery at m/z 1060.

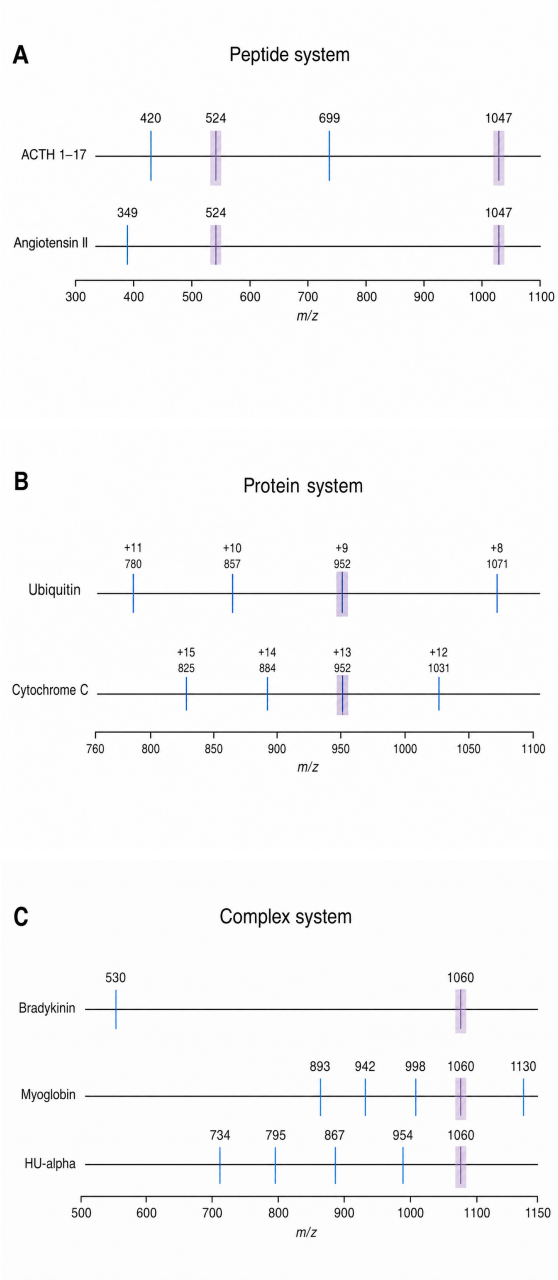


Figure 8: m/z-location summary for the three systems.

in the case of a charge-state route.

Compared to other peptide-centric DIA identification strategies through tandem MS, CSEP benefits from a narrower notion of contextual information. Tandem mass spectrometry identification based on peptides requires repeating windows, elution time patterns, or spectral libraries [11, 16]. Unlike those approaches, CSEP does not use any of them because the source of contextual information for this method is the charge-state organization of the precursor windows.

With regard to the available top-down proteoform identification software, CSEP performs a different function. Software products allow assigning proteoforms to a given list of product ions [13, 14]. Conversely, the function of CSEP lies in choosing whether the peaks in the mixed spectrum are to be retained, separated, or marked as ambiguous in assembling the product ion list.

3.6. Implications for MALDI MS/MS experimental design

It follows from the results above that the decision on selecting precursor ions should not be based exclusively on the cleanness of windows. In addition to using the cleanness, one may consider it more valuable to get more charge states rather than one spectrum in a single window. With regard to the peptide level, the windows common to the two peptides provided information to both peptides because they have been routed. With regard to proteins, several windows were needed to confirm their existence.

The optimal strategy can be derived from the coordinate summary in Fig. 8. The plan of acquisition, which fails to include m/z windows 524, 952, 1047, and/or 1060, is not efficient enough. There may be no way of telling where the information about the analyte is located - in this m/z windows or in the adjacent m/z windows or charge families. In this regard, acquiring and retaining precursor windows together with the charge assignment is crucial for the proper MALDI-MS/MS identification.

A direct application of the results of the study is the alternative reporting format. To complement database search score results, it is useful to report on the grounds for the database search score, i.e., what windows were shared and what peptide and protein windows got their scores boosted by the evidence partition.

4. Conclusion

A positive answer to the formulated question forms the core of this study: Can ambiguous evidence be used to assign a peptide/protein in the co-isolation of MALDI MS/MS measurements under the constraint of charge state context? The answer is 'yes' with important caveats. The shared window is not the evidence per se. Rather, the evidence arises in the course of recognizing the charge state context of this window. Hence, shared windows become recoverable only in the form of the analyte-specific path.

The interference-staged dataset serves to illustrate three degrees of difficulty. At the peptide level, ACTH 1–17 and angiotensin II demonstrated improvement from 71 to 153 and 25 to 66 respectively, indicating that shared peptide windows at m/z 524 and 1047 are reliable in routing even in the presence of minor channel products. The increase of ubiquitin from 42 to 288 and cytochrome C from 52 to 225 indicates that shared intact protein windows are supported by a charge state context. At the highest level, bradykinin increased from 3 to 56, myoglobin, up to 149, and HU-alpha, from 55 to 254. These findings indicate that shared windows at complex m/z 1060 can be recognized when the proper charge state evidence is available.

From this experimental study, one should derive the following conclusions about the utility of charge-state evidence. Although the primary function of this evidence type is boosting confidence in peptide assignment, its main function is serving as a criterion for filtering ambiguous evidence in a co-isolation MALDI MS/MS spectrum. From the standpoint of the experimental design, the results show the value of including charge state information to the MALDI protocol.

References

[1] Karas, M., & Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry, 60*(20), 2299-2301.

[2] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., & Whitehouse, C. M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science, 246*(4926), 64-71.

[3] Cramer, R., Pirkl, A., Hillenkamp, F., & Dreisewerd, K. (2013). Liquid AP-UV-MALDI enables stable ion yields of multiply charged peptide and protein ions for sensitive analysis by mass spectrometry. *Angewandte Chemie International Edition, 52*(8), 2364-2367.

[4] Ryumin, P., Brown, J., Morris, M., & Cramer, R. (2016). Investigation and optimization of parameters affecting the multiply charged ion yield in AP-MALDI MS. *Methods, 104*, 11-20.

[5] Houel, S., Abernathy, R., Renganathan, K., Meyer-Arendt, K., Ahn, N. G., & Old, W. M. (2010). Quantifying the impact of chimera MS/MS spectra on peptide identification in large-scale proteomics studies. *Journal of proteome research, 9*(8), 4152-4160.

[6] Michalski, A., Cox, J., & Mann, M. (2011). More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC–MS/MS. *Journal of proteome research, 10*(4), 1785-1793.

[7] Wang, J., Pérez-Santiago, J., Katz, J. E., Mallick, P., & Bandeira, N. (2010). Peptide identification from mixture tandem mass spectra. *Molecular & Cellular Proteomics, 9*(7), 1476-1485.

[8] Wang, J., Bourne, P. E., & Bandeira, N. (2014). MixGF: spectral probabilities for mixture spectra from more than one peptide. *Molecular & Cellular Proteomics, 13*(12), 3688-3697.

[9] Shteynberg, D., Mendoza, L., Hoopmann, M. R., Sun, Z., Schmidt, F., Deutsch, E. W., & Moritz, R. L. (2015). reSpect: software for identification of high and low abundance ion species in chimeric tandem mass spectra. *Journal of the American Society for Mass Spectrometry, 26*(11), 1837-1847.

[10] Tsou, C. C., Avtonomov, D., Larsen, B., Tucholska, M., Choi, H., Gingras, A. C., & Nesvizhskii, A. I. (2015). DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nature methods, 12*(3), 258-264.

[11] Yu, F., Teo, G. C., Kong, A. T., Fröhlich, K., Li, G. X., Demichev, V., & Nesvizhskii, A. I. (2023). Analysis of DIA proteomics data using MSFragger-DIA and FragPipe computational platform. *Nature communications, 14*(1), 4154.

[12] Lou, R., & Shui, W. (2024). Acquisition and analysis of DIA-based proteomic data: a comprehensive survey in 2023. *Molecular & Cellular Proteomics, 23*(2), 100712.

[13] Kou, Q., Xun, L., & Liu, X. (2016). TopPIC: a software tool for top-down mass spectrometry-based proteoform identification and characterization. *Bioinformatics, 32*(22), 3495-3497.

[14] Melby, J. A., Roberts, D. S., Larson, E. J., Brown, K. A., Bayne, E. F., Jin, S., & Ge, Y. (2021). Novel strategies to address the challenges in top-down proteomics. *Journal of the American Society for Mass Spectrometry, 32*(6), 1278-1294.

[15] Adair, L. R., Jones, I., & Cramer, R. (2024). Utilizing precursor ion connectivity of different charge states to improve peptide and protein identification in MS/MS analysis. *Analytical Chemistry, 96*(3), 985-990.

[16] Gillet, L. C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., ... & Aebersold, R. (2012). Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Molecular & Cellular Proteomics, 11*(6).