Novelty Assessment Report

Paper: CausalNovo: Advancing De Novo Peptide Sequencing via a Causality-Informed Framework
PDF URL: https://openreview.net/pdf?id=55GsILa8c9
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-07

Abstract

\emph{De novo} peptide sequencing is a foundational computational technique in proteomics, which is critical for discovering and characterizing novel peptides and proteins within complex biological systems. To predict peptide sequences directly from tandem mass spectra, mainstream deep learning approaches aim to model the relationship between mass spectra and corresponding peptides. However, these models face significant challenges, particularly under noisy conditions. These deep learning models often capture superficial correlations within noisy spectral data, failing to identify the underlying causal mechanisms that link true signal fragment ions to peptide sequences. Consequently, these models tend to learn spurious associations that cannot generalize in practice, where noise peaks are prone to change due to different co-elutions or chemical contaminants. To tackle this, we introduce CausalNovo, a model-agnostic framework designed to learn the causal representations of mass spectra in peptide sequencing models by focusing on signal fragment ions. Specifically, grounded in two practical and general principles, independence and sufficiency, CausalNovo employs causal interventions and information-theoretic objectives to disentangle causal representations from spurious noise peaks. Extensive experiments on three public datasets show that CausalNovo effectively generalizes across varying Noise Signal Ratios (NSR) and remains relatively stable against non-causal peak changes. Consequently, CausalNovo yields consistent and significant performance gains of up to 10\% in amino acid, peptide, and PTM-level performance. Code is available at https://anonymous.4open.science/r/CausalNovo-C134.

Core Task Landscape

This paper addresses: **de novo peptide sequencing from tandem mass spectra**

A total of **50 papers** were analyzed and organized into a taxonomy with **16 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Deep Learning-Based Sequencing Methods**
- **Traditional Algorithmic Sequencing Approaches**
- **Data-Independent Acquisition Sequencing**
- **Hybrid and Database-Assisted Sequencing**
- **Specialized Sequencing Contexts and Applications**
- **Methodological Foundations and Reviews**
- **Supporting Technologies and Enhancements**
- **Specialized Experimental and Implementation Studies**

Complete Taxonomy Tree

- de novo peptide sequencing from tandem mass spectra Survey Taxonomy
- Deep Learning-Based Sequencing Methods
  - Transformer and Attention-Based Models
  - Bidirectional and Encoder-Decoder Architectures (5 papers)
    - [2] Bidirectional de novo peptide sequencing using a transformer model (Sangjeong Lee, 2024) View paper
    - [3] PowerNovo: de novo peptide sequencing via tandem mass spectrometry using an ensemble of transformer and BERT models (D. V. Petrovskiy, 2024) View paper
    - [8] De novo mass spectrometry peptide sequencing with a transformer model (Melih Yilmaz, 2022) View paper
    - [9] BERT model for de novo protein sequencing using tandem mass spectra (D.V. Petrovsky, 2024) View paper
    - [30] Sequence-to-sequence translation from mass spectra to peptides with a transformer model (Bittremieux, 2023) View paper
  - Causality-Informed and Robust Learning Frameworks ★ (1 papers)
    - [0] CausalNovo: Advancing De Novo Peptide Sequencing via a Causality-Informed Framework (Anon et al., 2026) View paper
  - Convolutional and Recurrent Neural Networks (2 papers)
  - [5] Accurate de novo peptide sequencing using fully convolutional neural networks (Kaiyuan Liu, 2023) View paper
  - [41] De novo peptide sequencing by deep learning (Hieu Tran, 2017) View paper
  - Post-Translational Modification-Aware Models (1 papers)
  - [25] Modanovo: A Unified Model for Post-Translational Modification-Aware de Novo Sequencing Using Experimental Spectra from In Vivo and Synthetic Peptides (Daniela Klaproth-Andrade, 2025) View paper
- Traditional Algorithmic Sequencing Approaches
  - Graph-Based and Dynamic Programming Algorithms (3 papers)
  - [4] De Novo Peptide Sequencing via Tandem Mass Spectrometry (Theresa A. Addona, 1999) View paper
  - [31] MSNovo: a dynamic programming algorithm for de novo peptide sequencing via tandem mass spectrometry (Lijuan Mo, 2007) View paper
  - [45] Spectrum graph-based de-novo sequencing algorithm MaxNovo achieves high peptide identification rates in collisional dissociation MS/MS spectra (Petra Gutenbrunner, 2021) View paper

    - Tree Search and Hybrid Optimization Methods (1 papers)
    - [44] PepGo: a deep learning and tree search-based model for de novo peptide sequencing (Yuqi Chang, 2025) View paper
- Data-Independent Acquisition Sequencing (3 papers)
    - [28] Transformer-Based De Novo Peptide Sequencing for Data-Independent Acquisition Mass Spectrometry (Guo Xuan, 2023) View paper
    - [35] A transformer model for de novo sequencing of data-independent acquisition mass spectrometry data. (Justin Sanders, 2025) View paper
    - [42] BiATNovo: An Attention-based Bidirectional De Novo Sequencing Framework for Data-Independent-Acquisition Mass Spectrometry (Shu Yang, 2023) View paper
- Hybrid and Database-Assisted Sequencing (3 papers)
    - [10] Sequence database searches via de novo peptide sequencing by tandem mass spectrometry (J. A. Taylor, 1997) View paper
    - [27] Searching sequence databases via De novo peptide sequencing by tandem mass spectrometry (Richard S. Johnson, 2002) View paper
    - [33] Spectral dictionaries: Integrating de novo peptide sequencing with database search of tandem mass spectra (Sangtae Kim, 2009) View paper
- Specialized Sequencing Contexts and Applications
    - Top-Down and Protein-Level Sequencing (2 papers)
    - [16] De novo sequencing of peptides from top-down tandem mass spectra (Kira Vyatkina, 2015) View paper
    - [49] De Novo Protein Sequencing by Combining Top-Down and Bottom-Up Tandem Mass Spectra (Liu Xw, 2014) View paper
    - Cross-Linking and Specialized Fragmentation (3 papers)
    - [17] ECL 3.0: a sensitive peptide identification tool for cross-linking mass spectrometry data analysis (Zhou Chen, 2023) View paper
    - [22] Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry (John E. P. Syka, 2004) View paper
    - [38] Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry (Christian K. Frese, 2012) View paper
    - Small Molecule and Non-Peptide Sequencing (1 papers)
    - [6] MSNovelist: de novo structure generation from mass spectra (Michael A. Stravs, 2022) View paper
- Methodological Foundations and Reviews (8 papers)
    - [1] Lessons in de novo peptide sequencing by tandem mass spectrometry (KF Medzihradszky, 2015) View paper
    - [7] Algorithms for de-novo sequencing of peptides by tandem mass spectrometry: A review. (Cheuk Chi A. Ng, 2023) View paper
    - [12] Algorithms for the de novo sequencing of peptides from tandem mass spectra (Jens Allmer, 2011) View paper
    - [15] De novo sequencing of proteins by mass spectrometry (R. Vitorino, 2020) View paper
    - [19] De Novo Sequencing of Peptides from Tandem Mass Spectra and Applications in Proteogenomics. (SavaÅ Takan, 2024) View paper
    - [26] De Novo Sequencing of Peptides from Tandem Mass Spectra and Application in Proteogenomics (Takan, 2023) View paper
    - [43] Peptide and protein de novo sequencing by mass spectrometry (K. Standing, 2003) View paper
    - [46] Tandem Mass Spectrometry for Peptide Sequencing (Satyendra Mishra, 2025) View paper
- Supporting Technologies and Enhancements
    - Precision Mass Spectrometry and Isotopic Labeling (3 papers)
    - [23] De novo peptide sequencing and identification with precision mass spectrometry (Mikhail M. Savitski, 2007) View paper
    - [36] An algorithm for peptide de novo sequencing from a group of SILAC labeled MS/MS spectra. (Fang Han, 2025) View paper
    - [48] Rapid 'de novo' peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/timeâofâflight mass spectrometer (A. Shevchenko, 1997) View paper
    - Machine Learning-Based Feature Prediction and Rescoring (3 papers)
    - [29] MSBooster: improving peptide identification rates using deep learning-based features (Yang, 2023) View paper
    - [34] De novo peptide sequencing with InstaNovo: Accurate, database-free peptide identification for large scale proteomics experiments (Kevin Eloff, 2023) View paper
    - [40] Reducing Peptide Sequence Bias in Quantitative Mass Spectrometry Data with Machine Learning (Ayse B. Dincer, 2022) View paper
    - Spectral Quality and Noise Analysis (2 papers)
    - [39] De novo peptide sequencing by two-dimensional fragment correlation mass spectrometry (Zhongqi Zhang, 2000) View paper
    - [47] The impact of noise and missing fragmentation cleavages on de novo peptide identification algorithms (Kevin McDonnell, 2022) View paper
- Specialized Experimental and Implementation Studies (10 papers)
    - [11] Algorithms for de novo peptide sequencing using tandem mass spectrometry (Bingwen Lu, 2004) View paper
    - [13] Real-time peptide identification from high-throughput mass-spectrometry data (Sumesh Kumar, 2021) View paper
    - [14] Implementation and uses of automated de novo peptide sequencing by tandem mass spectrometry (J. Alex Taylor, 2001) View paper
    - [18] Sensitive, high-throughput HLA-I and HLA-II immunopeptidomics using parallel accumulation-serial fragmentation mass spectrometry (Kshiti Meera Phulphagar, 2023) View paper
    - [20] How to deal with internal fragment ions? (Arthur Grimaud, 2025) View paper
    - [21] Optimization of Collagenase Proteomics for Improved Mass Spectrometry Imaging Peptide Identification. (Jade K. Macdonald, 2025) View paper
    - [24] DIAmeter: matching peptides to data-independent acquisition mass spectrometry data (Yang Young Lu, 2021) View paper
    - [32] Phosphopeptide fragmentation and analysis by mass spectrometry (Paul J. Boersema, 2009) View paper
    - [37] Mass Spectrometry of Peptides (M.-C Petit, 2024) View paper
    - [50] SequenceAssembler: A tool for protein sequence assembly from mass spectrometry data. (Celso Vitor A Q Calomeno, 2025) View paper

Narrative

Core task: de novo peptide sequencing from tandem mass spectra. The field has evolved from traditional algorithmic approaches—such as dynamic programming and graph-based methods exemplified by early works like DeNovo Tandem[4] and Database Searches[10]—toward modern deep learning-based sequencing methods that now dominate recent research. The taxonomy reflects this shift, with top-level branches spanning Deep Learning-Based Sequencing Methods, Traditional Algorithmic Sequencing Approaches, Data-Independent Acquisition Sequencing, Hybrid and Database-Assisted Sequencing, and several specialized contexts. Within the deep learning branch, transformer and attention-based models have become particularly prominent, leveraging architectures inspired by natural language

processing to decode complex spectral patterns. Representative works include Transformer DeNovo[8], InstaNovo[34], and Bidirectional Transformer[2], which demonstrate how sequence-to-sequence frameworks can effectively map mass spectra to peptide sequences. Meanwhile, traditional methods and hybrid approaches continue to provide foundational insights, and specialized branches address niche experimental settings such as top-down proteomics, cross-linking studies, and immunopeptidomics.

Recent attention has focused on improving model robustness, generalization, and interpretability within transformer-based architectures. PowerNovo[3] and Fully Convolutional[5] models explore different neural designs to handle noisy or incomplete spectra, while works like BERT Sequencing[9] adapt masked language modeling strategies to peptide prediction. CausalNovo[0] sits within the causality-informed and robust learning frameworks subgroup, emphasizing principled approaches to model training that account for causal relationships and distributional shifts in mass spectrometry data. Compared to nearby transformer models such as PowerNovo[3] or InstaNovo[34], CausalNovo[0] distinguishes itself by integrating causal reasoning to enhance reliability and reduce biases that can arise from spurious correlations in training data. This focus on robustness addresses a key challenge across the field: ensuring that deep learning models generalize well beyond the specific datasets and experimental conditions on which they were trained.

Related Works in Same Category

No sibling papers were found in the same taxonomy leaf. A taxonomy-subtopic-level comparison will be produced instead.

Taxonomy-Level Summary

Both subtopics represent advanced transformer-based approaches to de novo peptide sequencing that go beyond standard autoregressive models. The original leaf focuses on incorporating causal reasoning and robustness mechanisms to handle data quality issues like noisy spectra and spurious correlations. The sibling subtopic emphasizes architectural innovations in how sequences are predicted, specifically through bidirectional processing or encoder-decoder frameworks that can leverage context from both directions.

Similarities: - Both represent specialized transformer architectures beyond standard unidirectional models - Both aim to improve prediction accuracy in de novo peptide sequencing - Both address limitations of basic transformer approaches through architectural or methodological enhancements - Both likely handle the same input (tandem mass spectra) and output (peptide sequences)

Differences: - Original leaf focuses on handling data quality issues (noise, spurious correlations) through causal reasoning and robustness mechanisms, while sibling focuses on prediction architecture (bidirectional vs unidirectional) - Original leaf emphasizes the reasoning framework (causality-informed, robustness), while sibling emphasizes the structural design (encoder-decoder, bidirectional) - Original leaf's exclusion criterion is about absence of causality/robustness components, while sibling's exclusion is about prediction directionality - Original leaf addresses 'why' the model makes robust predictions (causal mechanisms), while sibling addresses 'how' sequences are generated (architectural flow)

Suggested Search Directions: - Investigate whether bidirectional architectures inherently provide robustness benefits that overlap with causality-informed approaches - Explore hybrid models that combine bidirectional prediction with explicit causal reasoning mechanisms - Examine whether encoder-decoder frameworks can incorporate causal graph structures or robustness constraints

Sibling Subtopics

- **Bidirectional and Encoder-Decoder Architectures** (leaves: 1, papers: 5)
- Scope: Models predicting peptide sequences bidirectionally or using encoder-decoder frameworks with transformers.
- Exclude: Unidirectional or single-direction prediction models belong in standard transformer approaches.

Contributions Analysis

Overall novelty summary. The paper introduces CausalNovo, a model-agnostic framework applying causal reasoning to de novo peptide sequencing from tandem mass spectra. Within the taxonomy, it occupies a newly defined leaf node labeled 'Causality-Informed and Robust Learning Frameworks' under the broader 'Transformer and Attention-Based Models' branch. Notably, this leaf contains only the original paper itself, with no sibling papers identified, suggesting this represents a relatively sparse and emerging research direction within the deep learning-based sequencing landscape.

The taxonomy tree reveals that CausalNovo sits within a well-populated parent branch of transformer and attention-based models, which includes neighboring leaves such as 'Bidirectional and Encoder-Decoder Architectures' containing five papers. These sibling directions focus on architectural innovations like bidirectional prediction and encoder-decoder frameworks, whereas CausalNovo's leaf explicitly targets causal reasoning and robustness mechanisms. The taxonomy's scope note clarifies that standard transformer models without explicit causality components belong elsewhere, positioning this work as a methodological departure from purely architectural advances toward principled handling of noisy spectra and spurious correlations.

Across three identified contributions—the CausalNovo framework, structural causal model formalization, and independence-sufficiency principles—the analysis examined twenty candidate papers total, with five, six, and nine candidates respectively. Critically, zero refutable pairs were found for any contribution, meaning that among the limited set of top-K semantic matches and citation expansions examined, no prior work was identified that clearly overlaps with or anticipates these specific causal intervention strategies. This suggests that within the examined scope, the causal framing and information-theoretic objectives appear distinct from existing transformer-based sequencing methods.

Based on the limited literature search covering twenty candidates, the work appears to introduce a novel angle within deep learning-based peptide sequencing by explicitly incorporating causal reasoning. However, the analysis does not claim exhaustive coverage of all relevant prior work in causality or robustness for mass spectrometry, and the absence of sibling papers in the taxonomy leaf may reflect either genuine novelty or incomplete taxonomy construction rather than definitive field-wide uniqueness.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: CausalNovo framework for de novo peptide sequencing

Description: The authors propose CausalNovo, a model-agnostic framework that applies causal principles to de novo peptide sequencing. The framework learns causal representations from mass spectra by distinguishing signal fragment ions from spurious noise peaks, improving robustness and generalization across different noise conditions.

This contribution was assessed against **5 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Towards automated scientific discovery: Knowledge representation and reasoning in cell signalling networks

URL: View paper

Brief Assessment

Automated Discovery[53] focuses on knowledge representation and reasoning in cell signalling networks using logic modelling, Bayesian networks, and probabilistic relational models for phosphoproteomics data. It does not address de novo peptide sequencing from mass spectra or causal representation learning for this task.

---

2. PepNovo: de novo peptide sequencing via probabilistic network modeling

URL: View paper

Brief Assessment

PepNovo[52] focuses on probabilistic network modeling for scoring peptide-spectrum matches using fragmentation rules and likelihood ratios, not on causal representation learning to distinguish signal from noise peaks as in the original paper.

---

3. Prediction of peptide mass spectral libraries with machine learning

URL: View paper

Brief Assessment

Spectral Libraries[51] focuses on predicting peptide mass spectral libraries using machine learning, not on causal representation learning for de novo sequencing from mass spectra. The candidate addresses spectral library prediction rather than the causal framework proposed in the original paper.

---

4. Distilling Non-Autoregressive Model Knowledge for Autoregressive De Novo Peptide Sequencing

URL: View paper

Brief Assessment

Distilling Knowledge[54] focuses on knowledge distillation between autoregressive and non-autoregressive models for peptide sequencing, not on causal representations or distinguishing signal fragment ions from spurious noise peaks as in CausalNovo.

---

5. CHARACTERIZATION AND DE NOVO SEQUENCING OF MULTI-CHARGE MS/MS SPECTRA

URL: View paper

Brief Assessment

Multi-Charge Spectra[55] focuses on characterizing and sequencing multi-charge MS/MS spectra, not on causal representation learning or causal principles for de novo peptide sequencing. The candidate does not address causal mechanisms, spurious correlations, or causal interventions.

Contribution 2: Structural Causal Model formalization for peptide sequencing

Description: The authors formalize de novo peptide sequencing using Structural Causal Models to explicitly represent causal relationships between mass spectra and peptide sequences. This formalization distinguishes causal factors from non-causal spurious correlations, providing a principled foundation for robust model design.

This contribution was assessed against **6 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Faster graphical model identification of tandem mass spectra using peptide word lattices

URL: View paper

Brief Assessment

Peptide Lattices[61] focuses on computational efficiency improvements using word lattices and dynamic Bayesian networks for peptide sequencing, not on Structural Causal Model formalization or distinguishing causal from spurious correlations in mass spectra.

---

2. Modeling peptide fragmentation with dynamic Bayesian networks for peptide identification

URL: View paper

Brief Assessment

Dynamic Bayesian[60] uses probabilistic graphical models (DBNs) to model peptide fragmentation chemistry, not Structural Causal Models (SCMs) for representing causal relationships between mass spectra and peptide sequences as formalized in the original paper.

---

3. Graphical Models for Peptide Identification of Tandem Mass Spectra

URL: View paper

Brief Assessment

Graphical Models[59] focuses on probabilistic graphical models (Bayesian networks) for peptide identification from tandem mass spectra, not Structural Causal Models. The candidate does not address causal relationships, interventions, or the distinction between causal and spurious correlations that are central to the original paper's SCM formalization.

---

4. â¦ screening of umami peptides from skipjack tuna (Katsuwonus pelamis) hydrolysates using EAD/CID based micro-UPLC-QTOF-MS and the molecular interaction with â¦

URL: View paper

Brief Assessment

Skipjack Umami[57] focuses on identifying umami peptides from tuna hydrolysates using mass spectrometry, not on developing causal models for de novo peptide sequencing from tandem mass spectra.

---

5. Peptidomics-based analysis and preparation of umami peptides from enzymatically digested chicken bone fluid

URL: View paper

Brief Assessment

Chicken Umami[56] focuses on enzymatic digestion and umami peptide preparation from chicken bone fluid, not on causal modeling or computational peptide sequencing from mass spectra.

---

6. Umami peptides screened based on peptidomics and virtual screening from Ruditapes philippinarum and Mactra veneriformis clams

URL: View paper

Brief Assessment

Clam Umami[58] focuses on screening umami peptides from clams using peptidomics and virtual screening methods, not on developing Structural Causal Models for de novo peptide sequencing from mass spectra.

Contribution 3: Independence and sufficiency principles with information-theoretic objectives

Description: The authors derive two fundamental principles—independence (ensuring representations are invariant to non-causal factors) and sufficiency (retaining predictive information)—and operationalize them through causal interventions and information-theoretic objectives. These principles guide the disentanglement of causal signal from noise in the latent representation space.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Identifying Independencies in Causal Graphs with Feedback

URL: View paper

Brief Assessment

Feedback Independencies[70] focuses on identifying conditional independence relationships in causal graphs with feedback using d-separation criteria for discrete variables. The original paper addresses disentangling causal representations from noise in peptide sequencing using independence and sufficiency principles with information-theoretic objectives, which is a different application domain and methodological approach.

2. Causal Discovery with Continuous Additive Noise Models

URL: View paper

Brief Assessment

Additive Noise[66] focuses on causal discovery using structural equation models with additive noise, employing independence testing between residuals and inputs. The original paper's independence and sufficiency principles are designed for disentangling causal representations in peptide sequencing through interventions and mutual information objectives, which differs fundamentally from the candidate's regression-based independence testing approach.

3. Kernel-Based Independence Tests for Causal Structure Learning on Functional Data

URL: View paper

Brief Assessment

Functional Data[69] focuses on independence testing for functional/continuous data using kernel-based methods (HSIC, d-HSIC, HSCIC) for causal structure learning. The original paper addresses independence and sufficiency principles for disentangling causal representations in peptide sequencing through causal interventions and information-theoretic objectives in latent space. These are fundamentally different problem domains and methodological approaches.

4. Learning Causally Disentangled Representations via the Principle of Independent Causal Mechanisms

URL: View paper

Brief Assessment

Causally Disentangled[67] focuses on learning disentangled causal representations via independent causal mechanisms in VAE frameworks, not on de novo peptide sequencing with causal interventions to separate signal from noise peaks in mass spectrometry data.

5. CÂ²DR: Robust Cross-Domain Recommendation based on Causal Disentanglement

URL: View paper

Brief Assessment

C²DR[71] focuses on cross-domain recommendation with causal disentanglement of domain-shared vs. domain-specific preferences, not on general independence/sufficiency principles for causal representation learning in noisy data contexts like peptide sequencing.

6. Learning Independent Causal Mechanisms

URL: View paper

Brief Assessment

Independent Mechanisms[65] focuses on learning independent causal mechanisms as autonomous modules in generative models for domain adaptation and transfer learning. CausalNovo applies independence and sufficiency principles specifically to de novo peptide sequencing with causal interventions on spectral data, representing a distinct application domain and technical approach.

7. Partial disentanglement for domain adaptation

URL: View paper

Brief Assessment

Partial Disentanglement[63] focuses on domain adaptation with partitioned latent spaces (invariant/changing components) rather than causal peptide sequencing. The independence principle in [63] ensures invariance across domains via domain-specific transformations, not causal interventions for noise disentanglement in mass spectrometry.

8. Cause-effect inference in location-scale noise models: Maximum likelihood vs. independence testing

URL: View paper

Brief Assessment

Location-Scale Inference[68] focuses on causal direction inference in bivariate location-scale noise models using independence testing versus maximum likelihood, not on disentangling causal representations from noise in deep learning frameworks for peptide sequencing.

9. On Causally Disentangled State Representation Learning for Reinforcement Learning based Recommender Systems

URL: View paper

Brief Assessment

Disentangled Recommender[64] focuses on reinforcement learning-based recommender systems using conditional mutual information to identify causal state variables, not on disentangling causal representations from noise in peptide sequencing through independence and sufficiency principles.

Appendix: Text Similarity Detection

Textual similarity detection checked 21 papers and found 3 similarity segment(s) across 1 paper(s).

The following **1 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. Disentangling Causal Substructures for Interpretable and Generalizable Drug Synergy Prediction

Detected in: Contribution: contribution_3

⚠ **Note**: This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] CausalNovo: Advancing De Novo Peptide Sequencing via a Causality-Informed Framework View paper
- [1] Lessons in de novo peptide sequencing by tandem mass spectrometry View paper
- [2] Bidirectional de novo peptide sequencing using a transformer model View paper
- [3] PowerNovo: de novo peptide sequencing via tandem mass spectrometry using an ensemble of transformer and BERT models View paper
- [4] De Novo Peptide Sequencing via Tandem Mass Spectrometry View paper
- [5] Accurate de novo peptide sequencing using fully convolutional neural networks View paper
- [6] MSNovelist: de novo structure generation from mass spectra View paper
- [7] Algorithms for de-novo sequencing of peptides by tandem mass spectrometry: A review. View paper
- [8] De novo mass spectrometry peptide sequencing with a transformer model View paper
- [9] BERT model for de novo protein sequencing using tandem mass spectra View paper
- [10] Sequence database searches via de novo peptide sequencing by tandem mass spectrometry View paper
- [11] Algorithms for de novo peptide sequencing using tandem mass spectrometry View paper
- [12] Algorithms for the de novo sequencing of peptides from tandem mass spectra View paper
- [13] Real-time peptide identification from high-throughput mass-spectrometry data View paper
- [14] Implementation and uses of automated de novo peptide sequencing by tandem mass spectrometry View paper
- [15] De novo sequencing of proteins by mass spectrometry View paper
- [16] De novo sequencing of peptides from top-down tandem mass spectra View paper
- [17] ECL 3.0: a sensitive peptide identification tool for cross-linking mass spectrometry data analysis View paper
- [18] Sensitive, high-throughput HLA-I and HLA-II immunopeptidomics using parallel accumulation-serial fragmentation mass spectrometry View paper
- [19] De Novo Sequencing of Peptides from Tandem Mass Spectra and Applications in Proteogenomics. View paper
- [20] How to deal with internal fragment ions? View paper
- [21] Optimization of Collagenase Proteomics for Improved Mass Spectrometry Imaging Peptide Identification. View paper
- [22] Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry View paper
- [23] De novo peptide sequencing and identification with precision mass spectrometry View paper
- [24] DIAmeter: matching peptides to data-independent acquisition mass spectrometry data View paper
- [25] Modanovo: A Unified Model for Post-Translational Modification-Aware de Novo Sequencing Using Experimental Spectra from In Vivo and Synthetic Peptides View paper
- [26] De Novo Sequencing of Peptides from Tandem Mass Spectra and Application in Proteogenomics View paper
- [27] Searching sequence databases via De novo peptide sequencing by tandem mass spectrometry View paper
- [28] Transformer-Based De Novo Peptide Sequencing for Data-Independent Acquisition Mass Spectrometry View paper
- [29] MSBooster: improving peptide identification rates using deep learning-based features View paper
- [30] Sequence-to-sequence translation from mass spectra to peptides with a transformer model View paper
- [31] MSNovo: a dynamic programming algorithm for de novo peptide sequencing via tandem mass spectrometry View paper
- [32] Phosphopeptide fragmentation and analysis by mass spectrometry View paper
- [33] Spectral dictionaries: Integrating de novo peptide sequencing with database search of tandem mass spectra View paper
- [34] De novo peptide sequencing with InstaNovo: Accurate, database-free peptide identification for large scale proteomics experiments View paper
- [35] A transformer model for de novo sequencing of data-independent acquisition mass spectrometry data. View paper
- [36] An algorithm for peptide de novo sequencing from a group of SILAC labeled MS/MS spectra. View paper
- [37] Mass Spectrometry of Peptides View paper
- [38] Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry View paper
- [39] De novo peptide sequencing by two-dimensional fragment correlation mass spectrometry View paper
- [40] Reducing Peptide Sequence Bias in Quantitative Mass Spectrometry Data with Machine Learning View paper
- [41] De novo peptide sequencing by deep learning View paper
- [42] BiATNovo: An Attention-based Bidirectional De Novo Sequencing Framework for Data-Independent-Acquisition Mass Spectrometry View paper
- [43] Peptide and protein de novo sequencing by mass spectrometry View paper
- [44] PepGo: a deep learning and tree search-based model for de novo peptide sequencing View paper
- [45] Spectrum graph-based de-novo sequencing algorithm MaxNovo achieves high peptide identification rates in collisional dissociation MS/MS spectra View paper
- [46] Tandem Mass Spectrometry for Peptide Sequencing View paper
- [47] The impact of noise and missing fragmentation cleavages on de novo peptide identification algorithms View paper
- [48] Rapid 'de novo' peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/timeâ□□ofâ□□flight mass spectrometer View paper
- [49] De Novo Protein Sequencing by Combining Top-Down and Bottom-Up Tandem Mass Spectra View paper
- [50] SequenceAssembler: A tool for protein sequence assembly from mass spectrometry data. View paper
- [51] Prediction of peptide mass spectral libraries with machine learning View paper
- [52] PepNovo: de novo peptide sequencing via probabilistic network modeling View paper
- [53] Towards automated scientific discovery: Knowledge representation and reasoning in cell signalling networks View paper
- [54] Distilling Non-Autoregressive Model Knowledge for Autoregressive De Novo Peptide Sequencing View paper
- [55] CHARACTERIZATION AND DE NOVO SEQUENCING OF MULTI-CHARGE MS/MS SPECTRA View paper
- [56] Peptidomics-based analysis and preparation of umami peptides from enzymatically digested chicken bone fluid View paper
- [57] â□¦ screening of umami peptides from skipjack tuna (Katsuwonus pelamis) hydrolysates using EAD/CID based micro-UPLC-QTOF-MS and the molecular interaction with â□¦ View paper
- [58] Umami peptides screened based on peptidomics and virtual screening from Ruditapes philippinarum and Mactra veneriformis clams View paper
- [59] Graphical Models for Peptide Identification of Tandem Mass Spectra View paper
- [60] Modeling peptide fragmentation with dynamic Bayesian networks for peptide identification View paper

- [61] Faster graphical model identification of tandem mass spectra using peptide word lattices View paper
- [62] Disentangling Causal Substructures for Interpretable and Generalizable Drug Synergy Prediction View paper
- [63] Partial disentanglement for domain adaptation View paper
- [64] On Causally Disentangled State Representation Learning for Reinforcement Learning based Recommender Systems View paper
- [65] Learning Independent Causal Mechanisms View paper
- [66] Causal Discovery with Continuous Additive Noise Models View paper
- [67] Learning Causally Disentangled Representations via the Principle of Independent Causal Mechanisms View paper
- [68] Cause-effect inference in location-scale noise models: Maximum likelihood vs. independence testing View paper
- [69] Kernel-Based Independence Tests for Causal Structure Learning on Functional Data View paper
- [70] Identifying Independencies in Causal Graphs with Feedback View paper
- [71] CÂ²DR: Robust Cross-Domain Recommendation based on Causal Disentanglement View paper