Rapid Communications in Mass Spectrometry

WILEY

Rapid Communications in Mass Spectrometry

RESEARCH ARTICLE OPEN ACCESS

Pepsin Digestion for Proteomic Studies of the Human Hair Shaft

Darian H. Yee[1] | Rustam Mukhtarov[2] | Aashi D. Sharma[2] | Chen Lin[2] | Bingyun Sun[2]

[1]Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada | [2]Department of Chemistry, Simon Fraser University, Burnaby, British Columbia, Canada

Correspondence: Bingyun Sun (bingyun_sun@sfu.ca)

Received: 21 July 2025 | **Revised:** 28 November 2025 | **Accepted:** 10 December 2025

Keywords: hair proteome | human hair shaft | keratin-associated proteins | pepsin digestion | stoichiometry of type I and type II cuticular keratins | trypsin digestion

ABSTRACT

Rationale: Human hair shafts have received increased research interest owing to their easy accessibility and potential as a window for human health. Because the most abundant component in hair is protein, proteomics is a promising tool for studying the molecular composition of hair shafts. As one of the most sophisticated biomaterials, hair shafts also possess unique structures, particularly keratin intermediate filaments, posing challenges for proteomic sample processing. Previously, we discovered that incomplete trypsin proteolysis increased keratin sequence coverage but resulted in an abnormal stoichiometry between types I and II cuticular keratins (PMCID: 12130615).

Methods: In the present study, we explored the potential to re-examine the human hair proteome through pepsin proteolysis and evaluate whether the previously observed type II to type I keratin ratio was due to enzyme biases introduced particularly by trypsin digestion.

Results: After optimizing the pepsin digestion conditions, we not only confirmed that previous bias was indeed contributed by trypsin but also discovered that pepsin was more effective at identifying keratin-associated proteins, another main protein component than keratins in human hair shafts.

Conclusions: Spectral counting on trypsin-based proteomics has been widely used to study the stoichiometry of protein complexes. For the first time, we confirmed a large bias caused by the trypsin enzyme in spectral counting. We further demonstrated that the use of pepsin can effectively correct such bias. In addition, we discovered that pepsin digestion can better identify keratin-associated proteins than trypsin proteolysis, which offers another effective tool for studying the hair proteome.

1 | Introduction

Hair shafts have attracted unique research interest with a historical focus on wools, which are of great economic importance in the textile industry [1–4] and more recently a growing focus on human hair shafts in cosmetics, dermatology, and forensics [5–8]. Holistic interrogation of the structure and composition of human hair has been deployed for studies of metabolomics, lipidomics, genomics, proteomics, and phenomics [5]. Among these omic approaches, proteomics is the most promising method, as proteins are the major component of the hair shaft [5].

The ability to identify proteins in the hair shaft and detect changes in the hair proteome is anticipated opening many channels for health assessment and analysis, with possibilities to view the health history of individuals and even to determine the biological relationships between two people [9, 10]. In the past, hair follicles were intensively studied, in part owing to

their presence of stem cells, and less light has been shed on the hair shaft [11–14]. Unlike most human tissue samples, which are metabolically active and heavily influenced by the diurnal cycle and daily diet, the hair shaft is a stable reservoir of dead cells that have accumulated in the past and can be used to retrieve the health history of individuals [6]. With no active metabolism, the hair shaft preserves metabolites and proteins as it grows out of the follicle [6]. This information has been used to determine the influence of substances and for toxicology studies [15, 16]. The high quantity of keratins in the hair shaft has aided in the forensic identification of individuals [17].

Some difficulties that arise in the study of hair proteomes stem from the intrinsic properties of the hair shaft, such as the tough outer sheath of cuticular keratinocytes, which protects the internal cortex from solubilization and degradation and thus serves as an obstacle to protein extraction in proteomic studies [18–21]. The outer cuticle of hair consists of four sublayers (the epicuticle, A layer, exocuticle, and endocuticle), which present an armor plate of flat overlapping dead cells that are 160–714 nm thick [22] and protect the hair fiber from chemical and physiological weathering. Molecularly, the cuticle is known to contain a high abundance of cysteine residues compared with other layers of hair, which form dense cross links in addition to isopeptide bonds for the toughness of cuticles [23, 24]. The largest and most prominent layer of hair is the cortex, which contains melanin that gives the hair its color [15]. The cortex is made up of a collection of keratinocytes that are different in shape from keratinocytes in the cuticle. Cortical keratinocytes are long, spindle-shaped, and oriented along the axis of the hair. The medulla is a region at the center of the hair fibre that may be discontinuous or not present at all [25].

Each hair cortex cell contains 500–800 keratin intermediate filaments (KIFs) [26] which are organized into two families, i.e., the acidic keratin I family (K31–40) and the neutral or basic keratin II family (K81–86) [27]. Keratins are α-helical in shape, and a helix from each family will coil around each other to form a coiled-coil structure, which will then coil around another keratin coiled-coil in an antiparallel fashion to form a tetramer [28, 29] as shown in Figure 1. These tetramers are believed to then arrange head-to-tail (C-terminus to N-terminus) to form intermediate chains, which give hair the characteristic strength and durability [22, 30]. Keratin-associated proteins (KAPs) surround KIFs, forming an interfilamentous cross-linked matrix [31, 32] as shown in Figure 1 chiefly through disulfide linkages among other interactions, which create a strong hair shaft that is resistant to chemical treatment, UV exposure, environmental weathering, microbial attacks, and pollutants [18–21]. There are 23 KAP families [32], and certain KAPs are also thought to be responsible for assisting the assembly of KIFs, namely KAP8, which is a high glycine-tyrosine protein, and KAP11, a high-sulfur protein [32].

A better characterization of the hair shaft proteome directly affects its versatile and promising applications in human health and forensics. Previously, we reported that basic cuticular keratins possess dense trypsin cleavage sites, and complete digestion resulted in many short undetectable peptides, fewer detected peptides, and low sequence coverage [33]. This problem was overcome by incomplete digestion, which is typically unfavorable in proteomics sample preparation. Our study, however, demonstrated that incomplete trypsin digestion generated relatively longer peptides and achieved better sequence coverage of hair keratins [33]. In addition, both our

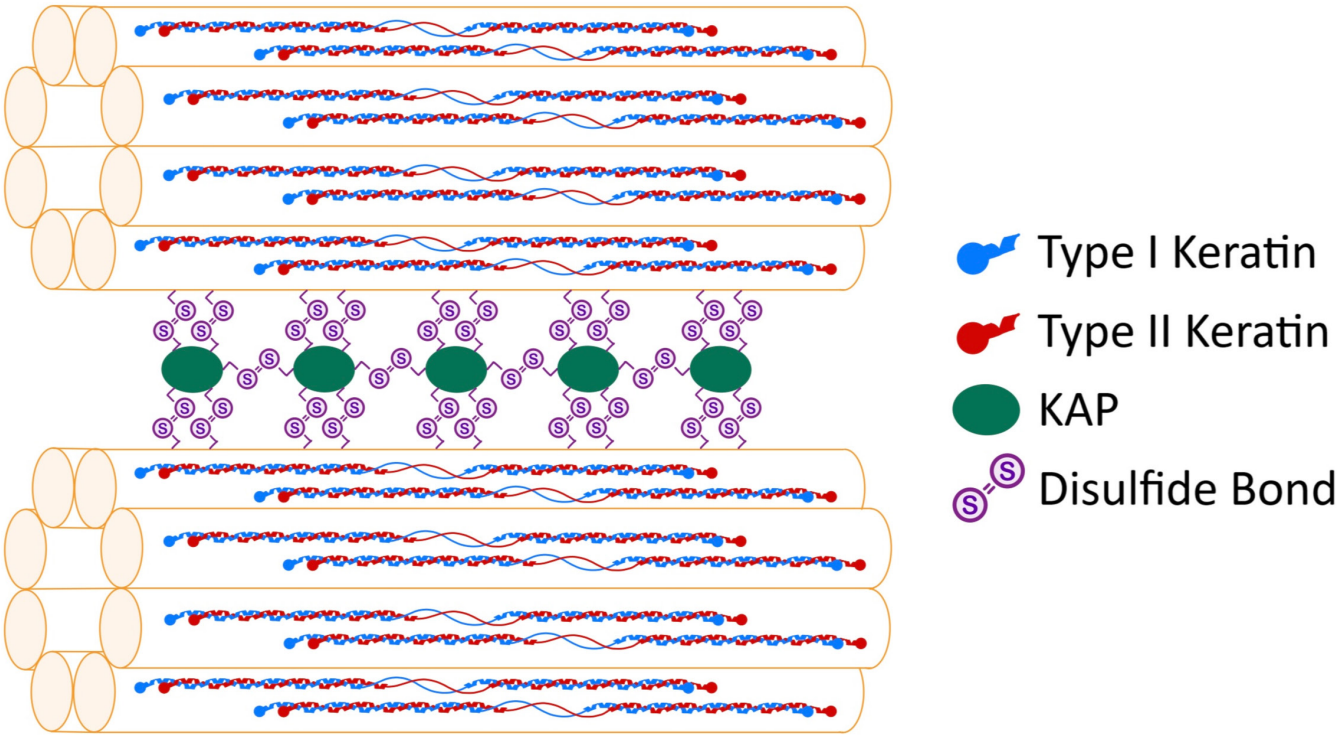

FIGURE 1 | Illustration of keratin intermediate filaments in human hair shafts.

identified hair proteome and previously published hair proteomes ubiquitously suggested that the ratio between acidic and basic keratins was not 1:1 as the literature has indicated, but rather was more than 2:1 [33]. Because acidic and basic keratins carry drastically different amounts of K and R, which are trypsin cleavage sites, we hypothesize that the observed ratio from a semi-quantitative spectral counting approach in proteomics may be influenced by the bias of trypsin cleavage sites among these proteins.

In addition to trypsin, many enzymes are capable of digesting proteins. In particular, pepsin has been used to digest keratins [34, 35] and has been widely used in the hide industry [36]. Unlike trypsin, which cleaves basic residues, pepsin cleaves hydrophobic and aromatic residues [37]. In the present study, we investigated how the proteomic quantitation generated by pepsin digestion can help clarify the discrepancy in the ratios of basic and acidic keratins between those obtained by tryptic proteomics and those obtained in the literature as mentioned above.

2 | Materials and Methods

2.1 | Materials

The mini-PROTEAN SDS-PAGE gel system was purchased from Bio-Rad Laboratories Inc. (Hercules, California, United States). Tris(2-carboxyethyl)phosphine (TCEP), beta-mercaptoethanol, dithiothreitol (DTT), sodium dodecyl sulfate (SDS), ethylenediaminetetraacetic acid (EDTA), iodoacetamide (IAA), and bovine serum albumin (BSA) were obtained from Sigma-Aldrich (St. Louis, Missouri, United States). Porcine pepsin was purchased from Promega Co. (Madison, Wisconsin, United States). The MCX cartridges were purchased from Waters Co. (Milford, Massachusetts, United States). LC-MS/MS instruments including an Easy nLC1000, a Q-Exactive HF Orbitrap, and a nano-EASY spray source were all purchased from Thermo Fisher Scientific (Mississauga, Ontario, Canada), and the other chemicals were also purchased from Thermo Fisher Scientific (Mississauga, Ontario, Canada).

2.2 | Protein Extraction

The sampling protocol followed a published procedure [33] that was approved by the Simon Fraser Ethics Board (application number: 30001840). Scalp hair samples were taken from various parts of the donor and trimmed 1 cm from the end to minimize potential contamination by the hair roots. Typically, 3 mg of the collected hair shaft was first sterilized with ethanol and homogenized using the Precelly 24, a lysis and homogenization device, in a lysis buffer consisting of 0.2 M NaOH, 1% SDS, 2% beta-mercaptoethanol, and 0.01 M EDTA in ddH_2O as detailed previously [33].

The obtained hair lysate was centrifuged at 12k rpm for 10 min (Microcentrifuge Sorvall Legend Micro 21, Thermo Fisher Scientific, Mississauga, Ontario, Canada), and the supernatant was subjected to acetone precipitation. Specifically chilled acetone at −20°C was applied to the hair lysate supernatant at a ratio of 4:1, mixed well, and stored at −20°C overnight. The sample was subsequently centrifuged at 12k rpm for 10 min on the second day. The protein pellet was collected and re-dissolved in 300 μL of 0.2% SDS, and the resolved concentration of proteins was measured via a bicinchoninic acid (BCA) assay (Thermo Fisher Scientific, Mississauga, Ontario, Canada).

2.3 | Pepsin Digestion

The pepsin digestion conditions were optimized on BSA first and then on hair shaft proteins. BSA of 0.1 mg in 10 μL was dissolved in 46.5 μL of 10 mM Tris–HCl at pH 7.5 with 0.01 M TCEP and 0.5% SDS to form the denaturation solution. The BSA sample was boiled for 10 min in a water bath, IAA was added to the sample solution at a final concentration of 0.01 M, and the sample was incubated in the dark at room temperature for 30 min with end-to-end rotation. DTT of 10 mM was subsequently added to the BSA solution to quench the excess IAA at room temperature for 15 min with rotation. HCl of 1 M was added to the final BSA solution to acidify the solution to pH 1–2. Pepsin was added to the BSA sample at a 1:50 pepsin-to-protein ratio, and the mixture was incubated at 37°C for overnight. The reaction was terminated by heating the sample to 95°C for 10 min. Aliquots of the sample before and after digestion were taken for SDS-PAGE evaluation, and the remaining proteolysis product was subjected to a mixed mode solid-phase extraction using MCX columns from Waters Co. (cat. 186000252) for cleaning, which can effectively remove SDS. In particular, a 1-cc MCX column was primed with 1 mL of methanol (MeOH) followed by 1 mL of 0.1% trifluoroacetic acid (TFA) before loading the acidified sample (pH 2–3) by gravity. The column was subsequently washed with 1 mL of 0.1% TFA and then 1 mL of 80% acetonitrile in 0.1% TFA to remove residual SDS and other non-peptide chemicals. Finally, the peptides were eluted from the column by 1 mL of 5% NH_4OH in MeOH by gravity.

A similar protocol with minor modifications was applied to 0.1 mg of hair shaft proteins in 10 μL of 0.2% SDS, and this sample was mixed with 10 mM Tris–HCl at pH 7.5 and TCEP for a final volume of ~60 μL and 0.01 M TCEP. The hair sample was first boiled at 100°C for 15 min, and then cooled to room temperature before 48 mg of urea powder was introduced. The sample was incubated at 37°C for 30 min before 96 μL of 0.5 M IAA was added, after which the sample was incubated in the dark at room temperature for 30 min with end-to-end rotation. Subsequently, 48 μL of 1 M DTT was added to quench the IAA by incubating at room temperature for 15 min with end-to-end rotation. The sample was diluted 10 times with ddH_2O, and 1 M HCl was used to acidify the hair sample to pH 1–2 before pepsin was added at 1:50 pepsin-to-protein ratio and incubated at 37°C for overnight. The digestion was terminated the next day by boiling the sample in the water bath for 15 min. Aliquots before and after pepsin digestion were taken for SDS-PAGE to evaluate the digestion efficiency. The remaining sample was subjected to MCX clean up as described above. The cleaned peptides were dried with a Thermo Fisher SpeedVac (model SPD121P-230, Mississauga, Ontario, Canada) and stored at −20°C.

2.4 | SDS-PAGE Evaluation of Digestion Efficiency

SDS-PAGE was employed to evaluate the initial digestion efficiency before LC-MS/MS proteomic analysis. Silver staining [38] was implemented for potential trace undigested proteins. ImageJ was used to qualitatively measure the stain intensities from images taken of the stained gels.

2.5 | LC-MS/MS Analysis

LC-MS/MS analysis was performed following a previously published procedure [33]. Specifically, the MCX cleaned samples were reconstituted in 0.1% formic acid (FA) prior to LC-MS/MS analysis. Separation by nanoLC was achieved using reversed-phase C18 trapping and analytical columns with 75 μm ID and 3 μm resin with 100 Å pores at lengths of 20 and 150 mm, respectively. A 60 min gradient of 2%–35% buffer B (0.1% FA in acetonitrile), as detailed previously [33], and a voltage of 2.3 kV was used for electrospray, and the ion transfer tube that guided ions into the MS was heated to 320°C. Data-dependent acquisition (DDA) mode was used for MS/MS analysis, and the top 10 most abundant precursor signals with charges greater than 1+ and fewer than 5+ in the MS1 scan were selected for higher energy collisional dissociation (HCD) under a dynamic exclusion of 10 s. The scan range for MS1 was set between 400–2000 m/z at a mass resolution of 120 000, and the resolution for the MS/MS scan was set at 30000. The automatic gain control (AGC) values for the MS1 and MS2 scans were 1×10^6 and 2×10^5, respectively. The default charge for the MS1 scans was 2+.

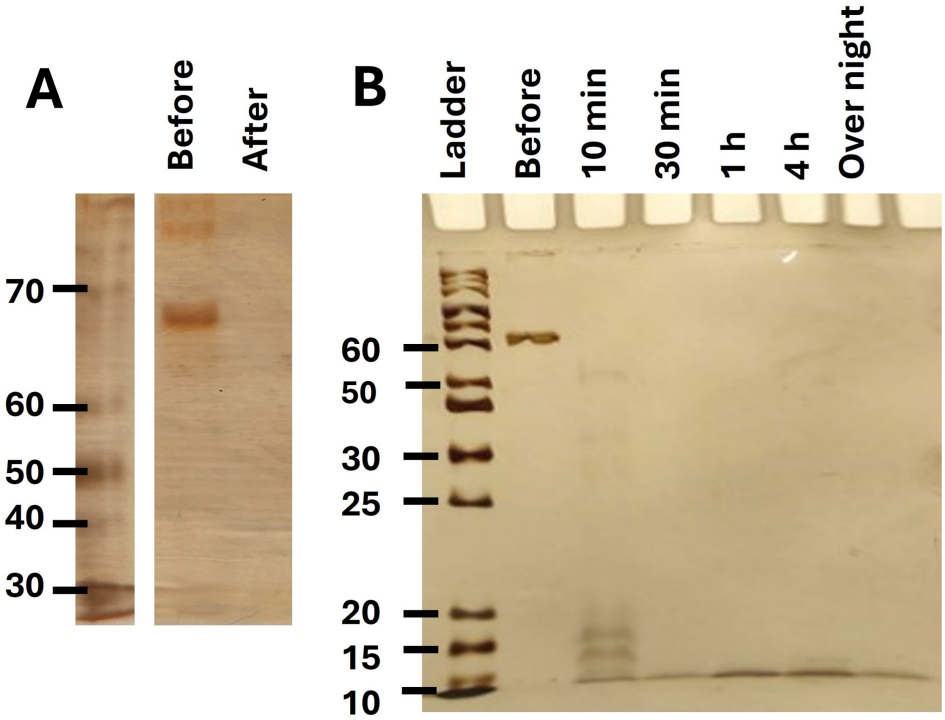

FIGURE 2 | SDS-PAGE results of pepsin digestion of bovine serum protein (BSA) under different incubation times.

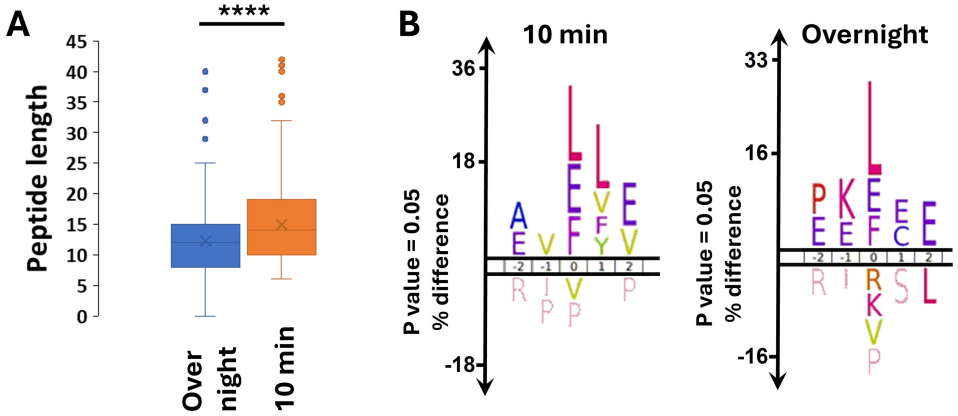

FIGURE 3 | Comparison of peptide length and cleavage sites after pepsin digestion with incubation for 10 min and overnight. (A) Boxplot of detected peptide length with the p value of the t test for 6E−6. (B) IceLogo results of the cleavage sites of the compared pepsin digests.

2.6 | Data Analysis

The raw files acquired by Xcalibur (Thermo Fisher Scientific, Mississauga, Ontario, Canada) from Q-Exactive HF were searched by Proteome Discoverer 2.1 (Thermo Fisher Scientific, Mississauga, Ontario, Canada) using Sequest HT. The human-reviewed UniProt entries were combined with a customized common contamination protein list that was used as the search database. The mass tolerance was 10 ppm for the monoisotopic precursors, and a mass tolerance of 0.02 Da was set for the monoisotopic fragments. No enzyme and a minimum of 6 amino acids were set as the search parameters. Modifications included a static modification of Cys by carbamidomethylation (+57.051 Da) and dynamic modifications of Met oxidation (+15.995 Da), Asp and Glu deamidation (+0.984 Da), and protein N-terminal acetylation (+42.011 Da). The false discovery rate (FDR) was obtained from the reversed sequences of the search database. A strict FDR cutoff of 0.01 and Percolator validation were used to obtain the final confident identifications.

A *t* test was performed using Microsoft Excel to evaluate the statistically significant differences among the means obtained for the compared values, in which the parameters were unpaired, two tails, and unequal variance.

Enzymatic cleavage site sequence analysis was performed by IceLogo [39]. In particular, 2–4 amino acid sequences up- and downstream of the C-termini of the detected peptides were subjected to analysis. The protein databases of *Homo sapiens* and *Bos taurus* were chosen as the background for the analysis of human hair proteins and BSA, respectively. A percentage difference and *p* value of 0.05 were chosen as the parameters for analysis.

3 | Results and Discussion

3.1 | Optimization of Pepsin Digestion

The incubation time of pepsin digestion was evaluated using BSA. Digestion was first evaluated by overnight incubation at 37°C, and complete digestion was observed as shown in Figure 2A. We then further evaluated the shorter duration of the digestion by dividing the overnight incubation into another

TABLE 1 | Comparison of peptides detected between trypsin and pepsin digestions.

	Trypsin digestion	Pepsin digestion	*t* test (*p*)
Peptide length	13.72	13.00	0.03
GRAVY score	−0.28	−0.39	0.03

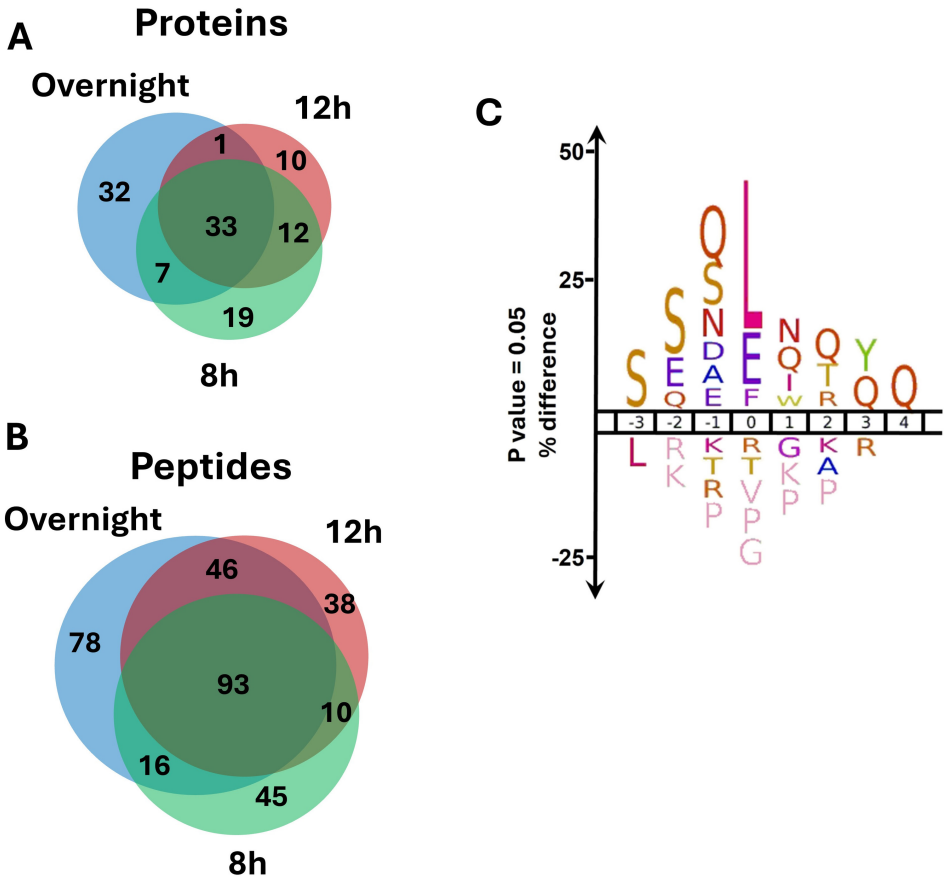

FIGURE 4 | Comparison of the hair proteome derived from pepsin digestion after different incubation times. (A) Comparison of detected proteins among the three timepoints. (B) Comparison of detected peptides among the three tested timepoints. (C) IceLogo analysis of commonly detected peptides across the three timepoints.

5 time points, i.e., 10 min, 30 min, 1 h, 4 h, and overnight incubation as shown in Figure 2B. All SDS-PAGE results revealed that pepsin rapidly digested BSA. We further used nano-LC-MS/MS to evaluate the detected BSA peptides after 10 min and overnight incubation as summarized in Table S1. The LC-MS/MS results revealed a decrease in peptide length with increasing incubation time as shown in Figure 3A, which was statistically significant with a p value of 6E−6. We also used IceLogo [39] to examine the cleavage specificity between the 10 min and overnight incubations as shown in Figure 3B. Both time points showed similar preferences for C-terminal Leu, Glu, and Phe, suggesting rapid action of the enzyme.

3.2 | Pepsin Digestion of Hair Proteins

We therefore evaluated the effects of pepsin digestion on hair proteins. Because hair proteins are notoriously resistant to proteolysis, we examined three relatively long digestion durations, i.e., 8, 12, and overnight. The results are summarized in Table S2, and in total, 114 human proteins were detected. We first compared the identified proteins in all three incubations, as shown in Figure 4A. Good agreement was observed among the three conditions, with overnight incubation resulting in the detection of most proteins, as anticipated. We further compared the peptides detected under these conditions, as shown in Figure 4B. The results resembled the protein results. A further Icelogo [39] analysis of the cleavage sites on the commonly detected peptides, as shown in Figure 4C resembled those obtained from BSA, with cleavage sites preferentially for Leu, Glu, and Phe. This preference is consistent with recent statistical modeling of the specificity of pepsin digestion [37].

3.3 | Comparison of the Hair Proteomes Obtained by Pepsin and Trypsin Digestion

To understand the difference in the impact on the detected human hair proteome, we compared the current pepsin hair proteome to the published trypsin hair proteome [33]. A minor increase of only 46 additional proteins was uniquely detected from pepsin digestion, which accounted for less than 10% increase from the hair proteome identified by trypsin digestion. A further comparison of their detected peptides is summarized in Table 1. The results revealed slightly but significantly shorter peptic peptides (average length: 13.01 amino acids) than tryptic peptides (average length: 13.72 amino acids) and slightly but significantly more hydrophilic peptic peptides (average GRAVY: −0.39) than tryptic peptides (average GRAVY: −0.28). Previously, a similar comparison between trypsin and pepsin digestion methods was carried out for pressurized online digestion, and the reported results also revealed a shorter peptide length and a shift to early retention time on a reversed-phase C18 column [40], which agreed with our observations in terms of both the peptide length and the GRAVY score. The accordance in results obtained from in-solution and online pressurized digestions suggests that both enzymes are robust with respect to their digestive conditions.

With the good agreement with previous studies, we further examined the stoichiometry between types I and II cuticular keratins based on the label-free quantitation offered by spectral counting [33, 41, 42]. Although far fewer proteins were detected by pepsin digestion than by trypsin digestion, all the hair keratin proteins were detected. Compared with the ~1:2.5 ratios between acid and basic cuticular keratins obtained by trypsin digestion [33], pepsin digestion resulted in a nearly 1:1 ratio between the same protein groups (Figure 5A), which is consistent with the known biological knowledge introduced in the Introduction section [28, 29]. Previously, the abnormal stoichiometry between the types I and II keratins observed in the hair proteome acquired from trypsin digestion was explained

TABLE 2 | Comparison of identified keratin-associated proteins via trypsin and pepsin digestion.

	Commonly detected	Uniquely detected by pepsin digestion
Keratin-associated proteins	KAP3-1, KAP3-3, KAP11-1, KAP19-5	KAP1-1, KAP1-3, KAP1-5, KAP3-2, KAP19-8

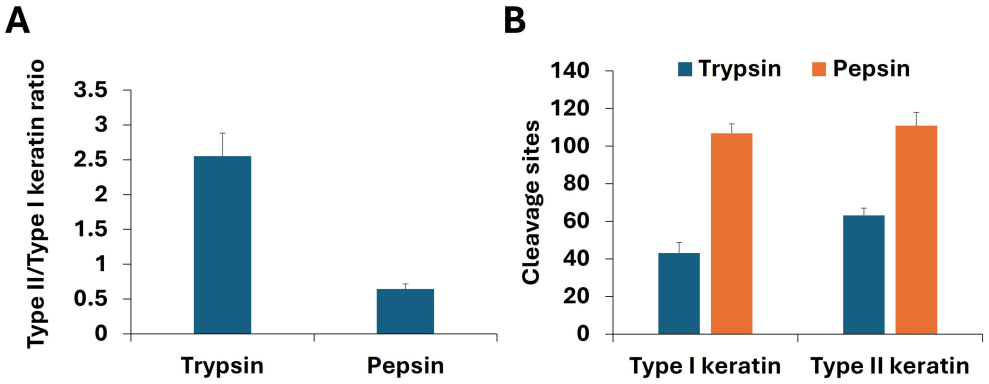

FIGURE 5 | Comparison of types I and II cuticular keratins. (A) Comparison of the ratios between types II and I keratins detected by trypsin and pepsin digestion. The error bars represent the standard deviation of the measured ratios. (B) Comparison of the theoretical cleavage sites of trypsin and pepsin among types I and II keratins. The error bars represent the standard deviation of the number of cleavage sites.

by the differential distribution of the basic trypsin cleavage sites (i.e., K and R) that are biased towards basic type II keratin detection [33]. We further analyzed the amino acid distribution of pepsin cleavage sites: L, F, and E obtained from Figures 3B and 4C among hair keratins and compared the results with those from trypsin cleavage sites. Figure 5B shows the results. Indeed, pepsin cleavage sites are less biased than trypsin cleavage sites between types I and II hair keratins, partially explaining the discrepancy in proteomic quantitation and biological values and further confirming the previous hypothesis that trypsin was likely the culprit of the observed biases in the quantitative proteomic results. Our study is also the first to demonstrate that spectral counting biases can be effectively corrected by using alternative enzymes such as pepsin when quantifying proteins with large differences in pI values.

In addition, among the unique proteins identified by pepsin digestion, several KAPs that were absent from the proteome derived by trypsin digestion were identified. As mentioned above, only a small fraction of proteins (< 10%) were uniquely contributed by pepsin digestion; however, pepsin digestion was able to result in all KAPs detected by trypsin with an additional five more KAPs, i.e., KAP1–1, KAP1–3, KAP1–5, KAP3–2, and KAP19–8 (Table 2). KAPs are known to form complexes with keratins according to 2D-PAGE proteomics [43], through interactions of disulfide bonds and isopeptide bonds among others; therefore, these proteins are difficult to identify. Among the five additionally identified KAPs,all of them carry twice the number of pepsin cleavage sites than the number of trypsin sites, which is likely to explain why pepsin can detect these KAPs more effectively. Even though trypsin proteolysis allowed the identification of a much larger hair proteome, pepsin proteolysis generated a much more comprehensive KAP identification. In proteomics, pepsin has been frequently implemented in proteomics and often in conjunction with trypsin to achieve optimal sequence coverage [44, 45]; here, we highlight another advantage of pepsin in proteomic analysis, i.e., its effectiveness in identifying KAPs.

4 | Conclusions

Here, we explored the use of pepsin to digest human hair for proteomic analysis. Our results showed that pepsin can digest model protein rapidly within minutes, but for hair samples, overnight digestion was more effective, with most protein and peptide identifications shown in Figure 4. In addition, our studies showed that pepsin digestion can effectively correct quantitation biases in acidic and basic cuticular keratins generated by commonly used trypsin digestion. The stoichiometry between types I and II hair keratins shown in Figure 5A from pepsin proteolysis is more adhered to the 1:1 biological knowledge than those from trypsin proteolysis. Finally, our results also indicated that pepsin is more effective than trypsin in identifying keratin-associated proteins in hair shafts, potentially because of its recognition of hydrophobic amino acids that are different from basic residues by trypsin. Although our results support the existing knowledge that a combination of trypsin and pepsin can generate a more comprehensive proteome, the effectiveness of pepsin in identifying keratin-associated proteins is novel to our study. We hope that our findings can promote future hair proteomics studies.

Author Contributions

B.S.: conceptualization, methodologies, supervision, and resources. D.H.Y., R.M., A.D.S., and C.L.: experiments. D.H.Y. and B.S.: writing–original draft preparation.

Acknowledgments

This work is supported by the National Sciences and Engineering Research Council of Canada (RGPIN06073), the Canada Foundation of Innovation, and the British Columbia Knowledge Development Fund.

Funding

This work is supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN06073), the Canada Foundation for Innovation, and the British Columbia Knowledge Development Fund.

Disclosure

All authors have reviewed and agreed to the submitted manuscript.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. A. M. Almeida, J. E. Plowman, D. P. Harland, et al., "Influence of Feed Restriction on the Wool Proteome: A Combined iTRAQ and Fiber Structural Study," *Journal of Proteomics* 103 (2014): 170–177.

2. X. Chen and J. W. S. Hearle, "Structural Hierarchy in Textile Materials: An Overview," in *Modelling and Predicting Textile Behaviour* (Elsevier, 2010), 3–40.

3. J. A. Maclaren and B. Milligan, "Wool Science: The Chemical Reactivity of the Wool Fibre1981".

4. J. Zhao, N. Liu, K. Liu, et al., "Identification of Genes and Proteins Associated With Anagen Wool Growth," *Animal Genetics* 48, no. 1 (2017): 67–79.

5. H. A. Adeola, J. C. Van Wyk, A. Arowolo, R. M. Ngwanya, K. Mkentane, and N. P. Khumalo, "Emerging Diagnostic and Therapeutic Potentials of Human Hair Proteomics," *Proteomics. Clinical Applications* 12, no. 2 (2018): 1700048, https://doi.org/10.1002/prca.201700048.

6. H. A. Adeola, J. C. Van Wyk, A. T. Arowolo, and N. P. Khumalo, "Human Hair as a Testing Substrate in the Era of Precision Medicine: Potential Role of Omics-Based Approaches," in *Keratin*, ed. M. Blumenberg (IntechOpen, 2018).

7. A. J. Grosvenor, S. Deb-Choudhury, P. G. Middlewood, et al., "The Physical and Chemical Disruption of Human Hair After Bleaching - Studies by Transmission Electron Microscopy and Redox Proteomics," *International Journal of Cosmetic Science* 40, no. 6 (2018): 536–548.

8. R. H. Rice, B. P. Durbin-Johnson, S. M. Mann, et al., "Corneocyte Proteomics: Applications to Skin Biology and Dermatology," *Experimental Dermatology* 27, no. 8 (2018): 931–938.

9. N. Karim, T. J. Plott, B. P. Durbin-Johnson, et al., "Elucidation of Familial Relationships Using Hair Shaft Proteomics," *Forensic Science International: Genetics* 54 (2021): 102564.

10. P. W. Wu, K. E. Mason, B. P. Durbin-Johnson, et al., "Proteomic Analysis of Hair Shafts From Monozygotic Twins: Expression Profiles and Genetically Variant Peptides," *Proteomics* 17, no. 13-14 (2017): 1600462, https://doi.org/10.1002/pmic.201600462.

11. S. M. Fan, C. F. Tsai, C. M. Yen, et al., "Inducing Hair Follicle Neogenesis With Secreted Proteins Enriched in Embryonic Skin," *Biomaterials* 167 (2018): 121–131.

12. T. Nakatsuji, Y. Shi, W. Zhu, et al., "Bioengineering a Humanized Acne Microenvironment Model: Proteomics Analysis of Host Responses to Propionibacterium Acnes Infection In Vivo," *Proteomics* 8, no. 16 (2008): 3406–3415.

13. L. Wang, W. Xu, L. Cao, et al., "Differential Expression of Proteins Associated With the Hair Follicle Cycle - Proteomics and Bioinformatics Analyses," *PLoS ONE* 11, no. 1 (2016): e0146791.

14. R. Williams, G. E. Westgate, A. D. Pawlus, S. K. Sikkink, and M. J. Thornton, "Age-Related Changes in Female Scalp Dermal Sheath and Dermal Fibroblasts: How the Hair Follicle Environment Impacts Hair Aging," *Journal of Investigative Dermatology* 141, no. 4S (2021): 1041–1051.

15. J. A. Hubbard, "Review on Toxicology Testing in Hair," *Journal of Applied Laboratory Medicine* 10 (2025): 983–1000.

16. J. J. Palamar and A. Salomone, "On the Challenges of Hair Testing to Detect Underreported Substance Use in Research Settings," *American Journal of Drug and Alcohol Abuse* 49, no. 1 (2023): 1–4.

17. Z. Liu, H. Simayijiang, Q. Wang, et al., "DNA and Protein Analyses of Hair in Forensic Genetics," *International Journal of Legal Medicine* 137, no. 3 (2023): 613–633.

18. N. R. Barthelemy, A. Bednarczyk, C. Schaeffer-Reiss, D. Jullien, A. Van Dorsselaer, and N. Cavusoglu, "Proteomic Tools for the Investigation of Human Hair Structural Proteins and Evidence of Weakness Sites on Hair Keratin Coil Segments," *Analytical Biochemistry* 421, no. 1 (2012): 43–55.

19. Z. C. Goecker, M. R. Salemi, N. Karim, B. S. Phinney, R. H. Rice, and G. J. Parker, "Optimal Processing for Proteomic Genotyping of Single Human Hairs," *Forensic Science International: Genetics* 47 (2020): 102314.

20. Y. J. Lee, R. H. Rice, and Y. M. Lee, "Proteome Analysis of Human Hair Shaft: From Protein Identification to Posttranslational Modification," *Molecular & Cellular Proteomics* 5, no. 5 (2006): 789–800.

21. S. Y. Wong, C. C. Lee, A. Ashrafzadeh, S. M. Junit, N. Abrahim, and O. H. Hashim, "A High-Yield Two-Hour Protocol for Extraction of Human Hair Shaft Proteins," *PLoS ONE* 11, no. 10 (2016): e0164993.

22. F. C. Yang, Y. Zhang, and M. C. Rheinstadter, "The Structure of People's Hair," *PeerJ* 2 (2014): e619.

23. R. D. B. Fraser, T. P. MacRae, L. G. Sparrow, and D. A. D. Parry, "Disulphide Bonding in Alpha-Keratin," *International Journal of Biological Macromolecules* 10 (1988): 106–112.

24. B. Wang, W. Yang, J. Mckittrick, and M. A. Meyers, "Keratin: Structure, Mechanical Properties, Occurrence in Biological Organisms, and Efforts at Bioinspiration," *Progress in Materials Science* 76 (2016): 229–318.

25. B. Buffoli, F. Rinaldi, M. Labanca, et al., "The Human Hair: From Anatomy to Physiology," *International Journal of Dermatology* 53, no. 3 (2014): 331–341.

26. R. Dawber, "Hair: Its Structure and Response to Cosmetic Preparations," *Clinics in Dermatology* 14, no. 1 (1996): 105–112.

27. R. Moll, M. Divo, and L. Langbein, "The Human Keratins: Biology and Pathology," *Histochemistry and Cell Biology* 129, no. 6 (2008): 705–733.

28. M. Er Rafik, J. Doucet, and F. Briki, "The Intermediate Filament Architecture as Determined by X-Ray Diffraction Modeling of Hard Alpha-Keratin," *Biophysical Journal* 86, no. 6 (2004): 3893–3904.

29. R. D. B. Fraser and D. A. D. Parry, "Structural Hierarchy of Trichocyte Keratin Intermediate Filaments," *Advances in Experimental Medicine and Biology* 1054 (2018): 57–70.

30. C. Popescu and H. Hocker, "Hair—The Most Sophisticated Biological Composite Material," *Chemical Society Reviews* 36, no. 8 (2007): 1282–1291.

31. D. A. Parry, T. A. Smith, M. A. Rogers, and J. Schweizer, "Human Hair Keratin-Associated Proteins: Sequence Regularities and Structural Implications," *Journal of Structural Biology* 155, no. 2 (2006): 361–369.

32. M. A. Rogers, L. Langbein, S. Praetzel-Wunder, H. Winter, and J. Schweizer, "Human Hair Keratin-Associated Proteins (KAPs)," *International Review of Cytology* 251 (2006): 209–263.

33. P. Kiatthitinan, D. H. Yee, H. M. Q. Tran, and B. Sun, "Advantages of Incomplete Digestion in Human Hair Shaft Proteomics, a Focus on Cuticular Keratins," *Rapid Communications in Mass Spectrometry* 39, no. 17 (2025): e10071.

34. S. Bjelland, A. Gildberg, and G. Volden, "Degradation of Human Epidermal Keratin by Fish Pepsin," *Archives of Dermatological Research* 280, no. 2 (1988): 119–123.

35. S. Bjelland, G. Volden, and J. Raa, "A Simple Assay for Determining Keratin and Collagen Degradation In Vitro," *Archives of Dermatological Research* 280, no. 1 (1988): 50–53.

36. M. Siegler, M. Deselnicu, V. Bratulescu, and A. Anghel, "Process of Preparing Pepsin for Bating Hides," (1993), United States.

37. O. Suwareh, D. Causeur, J. Jardin, et al., "Statistical Modeling of In Vitro Pepsin Specificity," *Food Chemistry* 362 (2021): 130098.

38. R. C. Switzer, III, C. R. Merril, and S. Shifrin, "A Highly Sensitive Silver Stain for Detecting Proteins and Peptides in Polyacrylamide Gels," *Analytical Biochemistry* 98, no. 1 (1979): 231–237.

39. N. Colaert, K. Helsens, L. Martens, J. Vandekerckhove, and K. Gevaert, "Improved Visualization of Protein Consensus Sequences by iceLogo," *Nature Methods* 6, no. 11 (2009): 786–787.

40. D. Lopez-Ferrer, K. Petritis, E. W. Robinson, et al., "Pressurized Pepsin Digestion in Proteomics: An Automatable Alternative to Trypsin for Integrated Top-Down Bottom-Up Proteomics," *Molecular & Cellular Proteomics* 10, no. 2 (2011): S1–S11.

41. H. Liu, R. G. Sadygov, and J. R. Yates, 3rd, "A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics," *Analytical Chemistry* 76, no. 14 (2004): 4193–4201.

42. B. Sun, L. Ma, X. Yan, et al., "N-Glycoproteome of E14.Tg2a Mouse Embryonic Stem Cells," *PLoS ONE* 8, no. 2 (2013): e55722.

43. R. Sinclair, M. J. Flagler, L. Jones, N. Rufaut, and M. G. Davis, "The Proteomic Profile of Hair Damage," *British J. Dermatology* 166 (2012): 27–32.

44. M. Golizeh and L. Sleno, "Optimized Proteomic Analysis of Rat Liver Microsomes Using Dual Enzyme Digestion With 2D-LC-MS/MS," *Journal of Proteomics* 82 (2013): 166–178.

45. S. Perchepied, N. Eskenazi, C. Giangrande, et al., "Development of Immobilized Enzyme Reactors for the Characterization of the Glycosylation Heterogeneity of a Protein," *Talanta* 206 (2020): 120171.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table S1:** Detected BSA peptides from pepsin digestion by mass spectrometry. **Table S2:** Detected hair proteins from pepsin digestion by mass spectrometry.