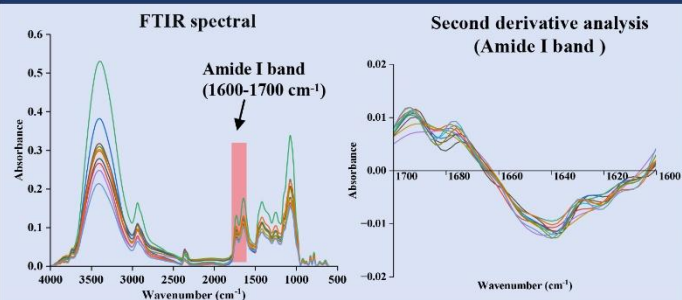


1 CP processing and FTIR data acquisition



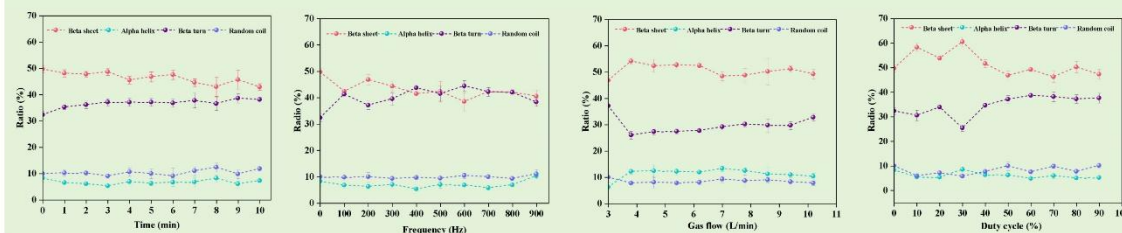
2 FTIR and spectral analysis



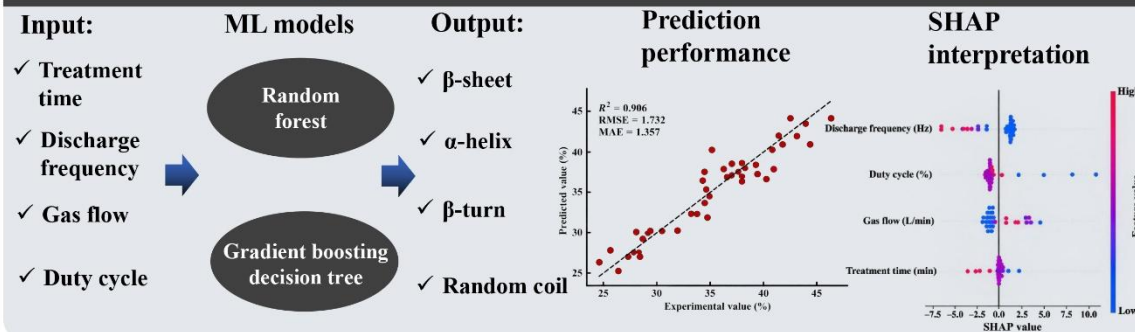
Peak deconvolution and secondary structure quantification

Peak No.	Secondary structure assignment	Wavenumber range (cm ⁻¹)	Fitted peak position (cm ⁻¹)	Peak area used for calculation	References
Peak 1	β -sheet	1615-1630	~1620	A1620	(Sadat & Joye, 2020)
Peak 2	β -sheet-related component	1630-1640	~1635	A1635	(Sadat & Joye, 2020)
Peak 3	Random coil	1640-1650	~1645	A1645	(Ye et al., 2007)
Peak 4	α -helix	1650-1660	~1655	A1655	(Goormaghtigh et al., 2006)
Peak 5	β -turn	1660-1680	~1670	A1670	(Sadat & Joye, 2020)
Peak 6	High-frequency amide I component associated with β -turn	1680-1700	~1690	A1690	(Sadat & Joye, 2020)

3 Secondary structure changes induced by CP parameters



4 ML prediction and interpretation



Highlights

- CP induced conformational rearrangement of kiwifruit proteins
- FTIR Amide I deconvolution quantified plasma-driven secondary structure transitions
- Discharge frequency and gas flow rate dominated structural modifications
- RF and GBDT models predicted CP-induced protein structural responses.
- SHAP analysis identified key parameters for targeted structural modulation.

1 **Cold plasma modulates the secondary structures of allergenic proteins**
2 **in postharvest kiwifruit: An FTIR and machine learning analysis**

3

4 **Shouqingyun Sun^a, Jilong Gao^a, Long Chen^{a,b,*}**

5 ^a *College of Mechanical and Electronic Engineering, Northwest A&F University,*
6 *Yangling, Shaanxi, 712100, China*

7 ^b *Department of Food Science and Technology, University of Nebraska-Lincoln, Lincoln,*
8 *NE, 68588, USA*

9

10

11

12

13

14

15

16

17

18

19

20 *Corresponding author. College of Mechanical and Electronic Engineering, Northwest
21 A&F University, Yangling, Shaanxi, 712100, China.

22 *E-mail address:* long.chen@nwafu.edu.cn (L. Chen).

23

Abstract

Postharvest kiwifruit contains allergenic proteins that can trigger immunoglobulin E-mediated immunoreactivity, highlighting the need for strategies to modulate protein structure. This study aimed to integrate Fourier transform infrared spectroscopy (FTIR) and machine learning (ML) to characterize cold plasma (CP)-induced changes in kiwifruit protein secondary structure. FTIR with Amide I band deconvolution was used to quantify β -sheet, α -helix, β -turn, and random coil contents under CP conditions. CP treatment reduced the β -sheet content from approximately 50% in the control to 40%-43% under intensified treatment, accompanied by an increase in β -turn content. Two ensemble ML algorithms, random forest (RF) and gradient boosting decision tree (GBDT), were used to predict protein secondary structures from CP processing parameters. Overall, both models accurately predicted β -turn ($R^2 > 0.90$) and α -helix ($R^2 \approx 0.89$), with slightly lower performance for β -sheet ($R^2 \approx 0.81$), whereas random coil prediction showed relatively lower accuracy. These results demonstrated the potential of ML to predict CP-induced protein structural changes, while SHAP analysis identified influential CP parameters, together supporting the selection of candidate processing conditions for targeted protein structural modulation. However, the effects of these structural changes on allergen immunoreactivity remain unclear. Future studies should integrate structural predictions with direct IgE-binding assays and other immunological assessments to clarify these relationships and guide immunoreactivity-oriented CP optimization in kiwifruit. Overall, this integrated FTIR-ML framework provided a methodological basis for targeted protein structural modulation during postharvest kiwifruit processing and offered a foundation for developing data-driven strategies for allergen management.

48 **Keywords:** Emerging food allergen; Modulation; Reactive oxygen species; Protein
49 structures; Amide I band

1. Introduction

Kiwifruit (*Actinidia deliciosa*) is widely recognized as a nutrient-rich fruit owing to its high contents of vitamins, minerals, dietary fiber, and various bioactive compounds, which contribute to its increasing popularity as a functional food worldwide (Wang et al., 2021b; Ao et al., 2025). However, despite its nutritional and health-promoting properties, postharvest kiwifruit has also been identified as an emerging food allergen that may induce adverse immune responses in susceptible individuals (Wang et al., 2021b; Deng et al., 2025a; Zhang et al., 2026). Consumption of kiwifruit has been associated with various allergic reactions, including oral allergy syndrome, gastrointestinal discomfort, skin irritation, and systemic allergic responses (Maddumage et al., 2013; Esteban-Sanz et al., 2025). These adverse reactions are primarily attributed to several allergenic proteins present in postharvest kiwifruit, such as Act d 1, Act d 2, Act d 5, and Act d 8, which can be recognized by allergen-specific immunoglobulin E (IgE) and subsequently trigger immune responses (Wang et al., 2019; Chen et al., 2026). Due to the increasing prevalence of food allergies and the lack of effective treatments other than allergen avoidance, reducing the allergenic potential of kiwifruit while preserving its nutritional and sensory qualities has become an important challenge. Immunoreactivity of proteins is not only determined by the presence of allergenic proteins but also strongly influenced by their structural characteristics, including folding patterns, conformational stability, and the accessibility of IgE-binding epitopes (Zhou et al., 2024b). The native conformation of allergenic proteins plays a critical role in maintaining specific epitopes that can be recognized by immune receptors. Structural modifications induced by food processing, such as unfolding, aggregation, oxidation, and rearrangement of secondary structures, may alter epitope exposure and subsequently affect IgE-binding capacity and immunoreactivity

([Rahaman et al., 2016](#); [Zhou et al., 2024b](#)). Therefore, monitoring secondary structural transitions provides valuable insights into the molecular mechanisms underlying allergen modification during food processing. Fourier-transform infrared spectroscopy (FTIR), particularly the Amide I band (1600-1700 cm⁻¹), has been widely applied for characterizing protein secondary structures due to its sensitivity to peptide backbone vibrations ([Kong & Yu, 2007](#)).

Cold plasma (CP), as an emerging non-thermal food processing technology, has attracted increasing attention for its potential application in modifying food components while minimizing thermal damage ([Ramazzina et al., 2015](#); [Zhou et al., 2024a](#); [Shah et al., 2025](#); [Yang et al., 2025](#)). Compared with conventional thermal processing, CP treatment operates under relatively mild temperature conditions while generating reactive oxygen and nitrogen species (RONS), charged particles, and ultraviolet photons ([Dharini et al., 2023](#); [Deng et al., 2025b](#)). These reactive species can interact with food macromolecules and induce structural and functional modifications ([Kaur et al., 2024](#)). Previous studies have demonstrated that CP-induced reactive species can promote protein oxidation, conformational rearrangement, aggregation, and changes in secondary structures, thereby affecting protein functionality and immunoreactivity ([Venkataratnam et al., 2020](#); [Liu et al., 2021](#); [Ng et al., 2021](#)). In particular, CP has shown potential in reducing the antigenicity of food allergens by modifying protein conformations and disrupting the accessibility of IgE-binding epitopes ([Hsieh et al., 2023](#); [Olatunde et al., 2023](#); [Hsieh & Ting, 2024](#)). Previous studies have employed CP for postharvest kiwifruit quality assessment ([Ramazzina et al., 2015](#)) and microbial decontamination ([Zhai et al., 2026](#)), yet its impacts on kiwifruit allergenic proteins remain largely unexplored.

In addition, the quantitative relationship between plasma processing parameters

and allergenic protein structural responses remains poorly understood. The coupled effects of discharge conditions, reactive species, and food matrix characteristics make it difficult to predict protein conformational changes from processing parameters alone, thereby limiting the rational optimization of CP conditions for targeted allergen modification. Conventional statistical approaches, including correlation analysis, analysis of variance, and regression-based methods, are useful for identifying individual associations and evaluating the effects of processing variables yet may be less effective in capturing complex nonlinear interactions among multiple parameters, such as treatment time, discharge frequency, duty cycle, and gas flow rate (Khan et al., 2022; Rashvand et al., 2025). Machine learning (ML) approaches provide an effective strategy for modeling complex nonlinear relationships and have been increasingly applied in food processing optimization and protein structure prediction due to their powerful feature extraction and predictive capabilities (Khan et al., 2022; Gao et al., 2025; Zhang et al., 2025; Tao et al., 2026). Ensemble ML algorithms, such as random forest (RF) and gradient boosting decision tree (GBDT), are particularly suitable for handling multidimensional processing data and identifying hidden relationships between processing conditions and structural responses (Breiman, 2001; Friedman, 2001). However, the “black-box” nature of ML models limits the interpretability of their predictions. Explainable artificial intelligence approaches, such as SHapley Additive exPlanations (SHAP), can therefore be used to quantify the relative contribution of individual processing parameters to model predictions and identify the processing variables that are most influential on the predicted protein structural responses (Lundberg & Lee, 2017; Parakh et al., 2025).

Therefore, this study aimed to characterize and predict CP-induced changes in the secondary structures of kiwifruit proteins and identify processing parameters that could

guide targeted structural modulation using an integrated FTIR and ML framework. The specific objectives were to: 1) quantify CP-induced secondary structural changes of CP-treated kiwifruit proteins using FTIR-based Amide I band deconvolution, 2) develop ML models to predict secondary structure responses from CP processing parameters, 3) use SHAP analysis to identify key parameters associated with specific secondary structure components and guide the selection of CP conditions for targeted structural modulation and subsequent immunoreactivity-oriented evaluation.

2. Materials and methods

2.1. Sample preparation

Fresh Xuxiang kiwifruits (*Actinidia deliciosa*) were purchased from a local market and stored at room temperature (approximately 25 °C) until reaching commercial ripeness (15-16 °Brix, total soluble solids). Xuxiang kiwifruit was selected because it is a commercially important cultivar widely cultivated in Shaanxi Province, a major kiwifruit-producing region in China. Commercial ripeness was defined as 15-16 °Brix based on previously reported physicochemical characteristics of commercially ripe kiwifruit (Zhang et al., 2020; Ao et al., 2026). Fruits with a uniform diameter of 50-60 mm and no visible defects were selected for subsequent experiments (Ramazzina et al., 2015). According to a previously reported preparation procedure (Wang et al., 2021a), the ripe kiwifruits were manually peeled, diced into small pieces, and homogenized with deionized water at a mass ratio of 1:1 using a high-speed juicer (TM-FSJ-01, Ten Moons Intelligent Technology Co., Ltd., Foshan, China) to prepare a homogeneous kiwi juice sample. The 1:1 fruit-to-water ratio was adopted to obtain a homogeneous and sufficiently fluid sample for uniform exposure to the dielectric barrier discharge (DBD) plasma treatment while maintaining a high proportion of the original fruit matrix. The freshly prepared kiwi juice was stored at -4 °C before CP treatment and processed

within 30 min to minimize physicochemical changes during storage.

2.2. CP processing

Fig. 1 illustrates the atmospheric-pressure DBD CP treatment system, which mainly consists of a plasma power supply (CTP-2000KP, Nanjing Suman Plasma Technology Co., Ltd., Nanjing, China), voltage regulator (TDGC2-1, Nanjing Suman Plasma Technology Co., Ltd., Nanjing, China), pulse controller (PC-07, Nanjing Suman Plasma Technology Co., Ltd., Nanjing, China), reactor, and electrode support. Based on the experimental design (Fig. 2), approximately 180 mL of fresh kiwi juice was placed in the reactor and subjected to CP treatment under the preset operating conditions. The optimization of CP treatment was carried out using a single-factor experimental design. The operating conditions were initially set at a treatment time of 5 min, discharge frequency of 200 Hz, duty cycle of 50%, and air flow rate of 3 L min⁻¹. In the single-factor experiments, one plasma processing parameter was varied while the other parameters were kept constant. Samples were coded according to the varied processing parameter: T-1 to T-10 for treatment time (1-10 min), F-100 to F-900 for discharge frequency (100-900 Hz), G-3.0 to G-10.2 for gas flow rate (3.0-10.2 L min⁻¹), and D-10 to D-90 for duty cycle (10-90%). The frequency and duty cycle ranges were determined by the adjustable operating limits of the plasma treatment system. The treatment time and gas flow rate were selected based on commonly reported conditions for food-related CP processing and the stable operating conditions of the present plasma system ([Warne et al., 2021](#); [Laroque et al., 2022](#); [Horta et al., 2026](#)). Within these ranges, the levels were set at regular intervals of 1 min for treatment time, 100 Hz for discharge frequency, 0.8 L min⁻¹ for gas flow rate, and 10% for duty cycle to systematically cover the practical processing window and provide sufficient variation for subsequent ML analyses. For each treatment condition, an equal volume of kiwi juice was transferred

into the plasma treatment chamber and processed under the corresponding experimental parameters. The untreated kiwifruit juice was used as the control. After CP treatment, all samples were immediately frozen at -20 °C and subsequently freeze-dried using a freeze dryer (LGJ-10C, Sihuan Keyi Technology Development Co., Ltd., Beijing, China). Freeze-drying was performed at -70 ± 2 °C under a vacuum pressure of 6 Pa for 60 h, following the procedure reported by [Yue et al. \(2025\)](#). The resulting freeze-dried powders were finely ground using a ceramic mortar and pestle (Aoqi Lopu Trading Co., Ltd., Tianjin, China), transferred into polyethylene bags (80 mm × 120 mm), sealed, and stored in a desiccator until FTIR analysis. Because freezing and dehydration can potentially affect protein conformation through ice-crystal formation and changes in protein-protein or protein-water interactions, the same freezing and freeze-drying procedure was applied to all samples, including the untreated control, to minimize processing-related differences among groups.

2.3. FTIR spectral data acquisition

Based on the method described by [Kumar et al. \(2023\)](#) with minor modifications, approximately 1.0 mg of freeze-dried kiwi powder was mixed thoroughly with 100 mg of dried potassium bromide (KBr), followed by compression into transparent tablets for spectral measurement. The spectral data were acquired using an FTIR spectrometer (Vertex 70, BRUKER Corporation, Germany), with a scanning range of 600-4000 cm^{-1} , a spectral resolution of 4 cm^{-1} , and 16 scans per sample ([Ding et al., 2024](#)).

2.4. FTIR spectral pre-processing and protein secondary structure analysis

In FTIR analysis, the Amide I band (1600-1700 cm^{-1}) typically presents as a broad and highly overlapped spectral contour, rendering the direct identification and quantification of individual structural components highly challenging ([Kong & Yu, 2007](#)). Therefore, appropriate mathematical preprocessing is an essential prerequisite

prior to spectral deconvolution. Initially, the raw FTIR spectra were subjected to Savitzky-Golay (SG) smoothing. This method utilized a moving window fitted with polynomial least squares to effectively eliminate high-frequency instrumental noise and random baseline fluctuations (Ozturk et al., 2026). Compared to conventional averaging methods, SG smoothing significantly improved the signal-to-noise ratio while maximally preserving the essential spectral features, including intrinsic peak shapes and center frequencies (Chen & Wang, 2019). Subsequently, a second-derivative mathematical transformation was applied to the SG-smoothed spectra. The primary rationale for applying the second derivative was its ability to enhance spectral resolution and reduce baseline effects. Specifically, the second derivative reduced spectral baseline drift and minimized interference from linear background variations, thereby facilitating the identification of overlapping absorption bands (Yu et al., 2025). More importantly, the second derivative effectively narrowed broad absorption bands, resolving heavily overlapping sub-peaks hidden within the broader Amide I band (Yang et al., 2015). In the second derivative spectrum, these hidden sub-peaks manifest as distinct local minima, which act as the precise initial center positions necessary for the subsequent curve-fitting process. Based on these identified peak coordinates, the Amide I band was deconvoluted and mathematically fitted into six individual sub-peaks utilizing the Voigt function. As shown in Table 1, the assignments of these six fitted peaks were as follows: sub-peaks located at 1615-1640 cm^{-1} corresponded to β -sheet, those at 1640-1650 cm^{-1} were attributed to random coil, those at 1650-1660 cm^{-1} represented α -helix, and those at 1660-1700 cm^{-1} were assigned to β -turn. As shown in Table 2, the relative content of each secondary structure was quantified by calculating the ratio of its integrated peak area to the total area of the entire Amide I band.

2.5. ML modeling for structural prediction

In food processing, the relationships between processing conditions and macromolecular structural responses can involve multiple interacting variables and nonlinear effects, making data-driven modeling useful for predicting structural changes under different processing conditions. In this study, RF and GBDT were selected as two complementary tree-based ensemble ML algorithms because both can model nonlinear relationships and interactions among processing variables without requiring a predefined functional form (An et al., 2024; Guo et al., 2026). RF uses an ensemble of decorrelated decision trees to improve prediction stability, whereas GBDT sequentially builds trees to reduce prediction errors. These characteristics make both algorithms suitable for the present structured dataset with multiple CP processing variables and limited sample size (Breiman, 2001; Friedman, 2001; Gao et al., 2025). In addition, the tree-based structure of both models allows subsequent SHAP analysis to quantify the contribution of individual processing variables to the predicted secondary structure responses (Lundberg & Lee, 2017). We established a regression modeling framework to predict the relative contents of four protein secondary structures (α -helix, β -sheet, β -turn, and random coil) from four CP processing variables: treatment time, discharge frequency, air flow rate, and duty cycle. RF and GBDT were implemented using the MultiOutputRegressor framework in Scikit-learn. Although a multi-output regression framework was used for implementation, each secondary structure component was predicted using an independent regression estimator. Therefore, four regression models were established for each algorithm. The dataset comprised 216 observations obtained from 36 experimental conditions, with six replicate measurements for each condition. The dataset was partitioned into training and testing subsets at an 8:2 ratio. This ratio was selected to provide sufficient observations for model training while retaining an independent subset for evaluating predictive performance on held-out observations.

Stratified sampling based on combined treatment-condition labels generated from the four CP parameters was applied during partitioning to maintain comparable distributions of processing conditions between the two subsets. A fixed random seed (random_state = 42) was used to ensure the reproducibility of data partitioning and stochastic model training. The value of 42 was used solely as a fixed seed and had no specific methodological significance. The RF model was configured with n_estimators = 300, max_depth = 8, min_samples_split = 4, and min_samples_leaf = 2. The GBDT model was configured with n_estimators = 250, learning_rate = 0.05, max_depth = 4, min_samples_split = 4, and min_samples_leaf = 2. These hyperparameters were predefined to balance nonlinear fitting capacity and model complexity while limiting the risk of overfitting given the available sample size. The number of estimators was set to provide sufficient ensemble capacity, whereas maximum tree depth and minimum sample constraints were used to control excessive tree growth. For GBDT, a learning rate of 0.05 combined with 250 estimators was used to enable gradual boosting and stable model fitting. Model performance was evaluated using the coefficient of determination (R^2), root mean square error ($RMSE$), and mean absolute error (MAE), which were calculated separately for each secondary structure component. The R^2 was calculated as (1):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

The $RMSE$ was calculated as (2):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

The MAE was calculated as (3):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

where y_i is the experimental value of the i -th observation, $\hat{y}_i$ is the predicted value of the i -th observation, $\bar{y}$ is the mean of the experimental values, and n is the number of observations.

Lower *RMSE* and *MAE* values indicate smaller prediction errors, whereas an R^2 value closer to 1 indicates better agreement between the predicted and experimental values. In addition to evaluation using the independent testing dataset, model robustness was assessed using repeated five-fold cross-validation with 10 repetitions, resulting in 50 validation iterations (Bergstra & Bengio, 2012). The cross-validation procedure was performed using the entire dataset.

2.6. Model interpretation via SHAP framework

Although RF and GBDT can effectively model complex nonlinear relationships, their inherent “black-box” nature limits the interpretability of the relationships between CP treatment parameters and changes in protein secondary structure. To overcome this interpretability limitation and unveil the underlying physicochemical interactions, the SHAP approach, rooted in cooperative game theory, was incorporated to demystify the trained regression models. Given that both RF and GBDT are tree-based ensemble models, the high-speed TreeSHAP algorithm was specifically utilized to compute the SHAP values for each input variable. In this study, SHAP values were calculated for the independent test set to quantify the marginal contribution of each processing parameter to the predicted relative content of each protein secondary structure component. SHAP summary plots were used to assess the relative importance and direction of individual plasma parameters across the test samples, whereas dependence plots were used to examine the relationships between individual parameters and model

predictions, including potential nonlinear or threshold-like patterns. This approach provided model-based interpretation of how CP treatment parameters were associated with changes in the secondary structure of kiwifruit proteins.

2.7. Data and statistical analysis

To ensure statistical robustness and reproducibility, three independent biological replicates were conducted, with two technical replicate measurements performed for each biological replicate. For FTIR analysis, the spectra and corresponding second-derivative spectra were generated using the mean values obtained from six independent measurements for each sample. The protein secondary structure contents estimated by spectral fitting were calculated from the six replicate measurements and were reported as mean $\pm$ standard deviation (SD). Data analyses were performed using Python 3.14.6 in PyCharm 2025.1.6.1. Numerical data processing, spectral preprocessing, and ML regression modeling were conducted using NumPy 2.4.6, SciPy 1.18.0, and Scikit-learn 1.9.0, respectively. ML prediction plots and SHAP beeswarm plots were generated using Matplotlib 3.11.0 and SHAP 0.52.0, respectively, while other graphical analyses and Pearson correlation analysis were performed using OriginPro 2024 (OriginLab Corporation, Northampton, MA, USA).

3. Results and discussion

3.1. FTIR spectral characteristics

The FTIR spectra of freeze-dried kiwifruit powders subjected to distinct CP conditions (treatment duration, discharge frequency, gas flow rate, and duty cycle) were presented in Fig. 3a-d, respectively. These spectral profiles were generally consistent with those reported by [Wang et al. \(2021a\)](#) and [Wang et al. \(2023\)](#). All samples exhibited characteristic absorption regions typical of fruit juice macromolecules. In the high-frequency region, a broad band at 3300-3400 cm^{-1} was assigned to O-H and N-H

stretching vibrations, reflecting overlapping signals from proteins, endogenous saccharides, and organic acids. Additionally, C-H stretching was observed at 2930-2850 cm^{-1} , followed by the distinct Amide I band (1600-1700 cm^{-1} , predominantly peptide C=O stretching) and Amide II band (1500-1560 cm^{-1} , N-H bending coupled with C-N stretching) (Barth, 2007). The persistence of the major peak positions, together with the absence of new characteristic absorption bands, suggested that no major changes in the chemical structure of the polypeptide backbone were detectable by FTIR, whereas the spatial conformation and non-covalent interactions of kiwifruit proteins were likely altered by CP treatment (Tolouie et al., 2018). Despite the relatively stable peak positions, absorbance intensities exhibited pronounced and non-monotonic variations under different processing conditions. Moderate treatment duration and discharge frequency increased the overall spectral intensity, possibly due to structural rearrangement and exposure of previously buried polar functional groups (Segat et al., 2015). With prolonged treatment or increased frequency, absorbance subsequently decreased, which may have resulted from further structural rearrangement and intermolecular reassociation that altered the accessibility of surface functional groups (Ekezie et al., 2017). Similar non-linear responses were observed with gas flow rate and duty cycle. Moderate gas flow may have favored the distribution and interaction of reactive plasma species with protein surfaces, whereas extreme flow rates may have reduced their residence time or intensified surface etching. The marked decrease in absorbance at high duty cycles (D-70 to D-90) may further indicate excessive energy input and possible thermal accumulation, which could promote oxidation and structural cross-linking (Tolouie et al., 2018). These structural alterations may be relevant to protein immunoreactivity because conformational changes can modify the spatial arrangement and accessibility of IgE-binding epitopes. Previous studies have shown

that CP treatment can alter protein structure and, in some cases, reduce IgE-binding capacity (Ekezie et al., 2019; Venkataratnam et al., 2020). However, this relationship is not necessarily linear, as unfolding may expose buried epitopes, whereas aggregation, cross-linking, or conformational rearrangement may mask or disrupt them. Therefore, the observed changes in α -helix, β -sheet, β -turn, and random coil contents provide structural evidence of CP-induced conformational rearrangement but cannot be regarded as direct evidence of reduced immunoreactivity. Because antibody binding and clinical allergenicity were not directly evaluated, the FTIR-derived changes should be considered indirect evidence of potential alterations in immunoreactivity. The broad and highly overlapped Amide I band also limits direct visual identification of subtle secondary structure changes. Therefore, second-derivative analysis and Amide I band deconvolution were further employed to quantitatively characterize these changes, which could subsequently be modeled using ML approaches (Yang et al., 2015).

3.2. Second-derivative analysis of the Amide I band

The second-derivative spectra obtained under different treatment times, discharge frequencies, gas flow rates, and duty cycles were presented in Fig. 4a-d. The broad Amide I peak was successfully resolved into several distinct minima in each case. While the amplitudes and slight wavenumber shift of these minima varied dynamically in response to different treatment times, discharge frequencies, gas flow rates, and duty cycles, their core positions consistently fell within well-defined structural domains. As described in Section 2.4, these resolved minima provided the empirical basis for identifying four distinct secondary structures: β -sheet (1615-1640 cm^{-1}), random coil (1640-1650 cm^{-1}), α -helix (1650-1660 cm^{-1}), and β -turn (1660-1700 cm^{-1}) (Goormaghtigh et al., 2006; Ye et al., 2007; Sadat & Joye, 2020). It is worth noting that the primary objective of this second-derivative analysis is not to directly evaluate the

macroscopic structural damage, but rather to establish strict physical and mathematical constraints for subsequent spectral processing (Kong & Yu, 2007). The exact peak numbers and center wavenumbers identified from these minima serve as indispensable initial parameters (fixed boundaries) for multi-peak deconvolution using Voigt curve fitting. By utilizing these second-derivative minima as fixed center constraints, arbitrary or over-parameterized fitting is effectively prevented (Byler & Susi, 1986).

3.3. Quantitative analysis of secondary structures

Quantitative analysis of the deconvoluted Amide I band revealed highly dynamic and parameter-dependent interconversions among protein secondary structures within the kiwifruit matrix. In the control sample, the secondary structures of proteins were predominantly stabilized by ordered β -sheets (approximately 50%) and flexible β -turns (approximately 32%), with minor contributions from α -helices and random coils. These findings were consistent with the results reported by Wang et al. (2021a) and Wang et al. (2023). As treatment time (Fig. 5a) and discharge frequency (Fig. 5b) increased, β -sheet content gradually decreased to approximately 43% and 40%, respectively, whereas β -turn content increased correspondingly. This transition may be associated with increased exposure of proteins to RONS during CP treatment. These reactive species may induce oxidative modifications of susceptible amino acid residues and alter the non-covalent interactions that stabilize β -sheet structures (Wang et al., 2024; Bayati et al., 2026). Such conformational rearrangements were potentially relevant to protein immunoreactivity because changes in secondary and higher-order structures can alter the spatial arrangement and accessibility of IgE-binding epitopes. This relationship was particularly relevant to kiwifruit allergens. Wang et al. (2021a) reported that ultrasound-induced structural changes in kiwifruit proteins were accompanied by a 50% reduction in the IgE-binding capacity of Act d 2. Furthermore, Wang et al. (2020) also showed

that electric-field-induced conformational changes in Act d 2 were associated with a 75.3% reduction in IgE binding. These findings indicated that processing-induced conformational rearrangement can modify the antibody-recognition properties of kiwifruit allergens. However, the direction of this relationship was not determined solely by changes in an individual secondary structure component. Therefore, the progressive decrease in β -sheet content and increase in β -turn content observed in the present study may reflect conformational rearrangement and alter epitope accessibility, yet cannot by themselves be interpreted as direct evidence of reduced IgE recognition. Instead, these structural changes provided indirect evidence that CP may modify the immunoreactive properties of kiwifruit proteins, which require further validation using IgE-binding assays. In contrast, gas flow rate (Fig. 5c) and duty cycle (Fig. 5d) produced pronounced non-monotonic changes in secondary structure, likely reflecting differences in reactive-species transport, residence time, and energy input ([Gautam et al., 2024](#); [Rout et al., 2025](#)). At a moderate duty cycle (30%), β -sheet content temporarily increased, whereas it decreased markedly at higher energy inputs (e.g., 90%), possibly indicating an initial structural reassociation followed by more extensive oxidative modification ([Li et al., 2024](#); [Panklai et al., 2025](#)). Similar parameter-dependent structural rearrangements have been associated with changes in the antibody-binding properties of kiwifruit allergenic proteins ([Wang et al., 2020](#); [Wang et al., 2021a](#)). Thus, CP-induced structural evolution likely reflected a balance between reversible reorganization and extensive protein modification, resulting in treatment-dependent changes in potential epitope accessibility. These nonlinear responses also demonstrated the practical value of ML modeling. By establishing quantitative correlations between CP operational parameters and protein secondary-structure dynamics, the ML models enable targeted screening of desirable processing windows

and accurate prediction of structural behavior under unexplored parameter combinations. This approach provides a data-driven strategy for guiding the targeted modulation of protein secondary structure while reducing experimental trial-and-error in CP process optimization. This predictive capability provided a quantitative basis for integrating structural descriptors with antibody-binding measurements, thereby facilitating the development of immunoreactivity-oriented CP processing strategies.

3.4. Pearson correlation analysis between processing parameters and protein secondary structures

Pearson correlation analysis was performed to evaluate the linear relationships between CP processing parameters and protein secondary structure, as shown in Fig. 6. Among the four processing parameters, treatment time and discharge frequency showed relatively stronger correlations with changes in protein secondary structure than those of gas flow rate and duty cycle. Discharge frequency exhibited the strongest correlations among the investigated parameters, with a moderate negative correlation with β -sheet content ($r = -0.51$) and a moderate positive correlation with β -turn content ($r = 0.47$). These results indicated that higher discharge frequency was associated with lower β -sheet content and higher β -turn content. However, Pearson correlation only characterizes linear associations and cannot capture potential nonlinear relationships between CP parameters and protein structural changes. This result agreed with the secondary structure analysis presented in Section 3.3 and was consistent with previous studies reporting that higher plasma discharge intensity accelerates protein structural rearrangement by increasing the generation of reactive species (Bayati et al., 2026). Gas flow rate displayed the strongest positive correlation with α -helix content ($r = 0.66$), while showing negative correlations with β -turn ($r = -0.51$) and random coil ($r = -0.20$). Unlike treatment time and frequency, gas flow rate exhibited only a weak correlation

with β -sheet content ($r = 0.24$), suggesting that it mainly regulated the transport and residence time of reactive species rather than directly determining the degree of protein oxidation. Similarly, duty cycle showed moderate correlations with β -sheet ($r = -0.35$), β -turn ($r = 0.27$), and random coil ($r = 0.30$), demonstrating that duty cycle-induced effects relied on plasma energy input variation and exhibited weaker linear associations with protein structural evolution. Overall, the Pearson correlation analysis demonstrated that CP processing parameters contributed differently to protein conformational changes. Treatment time and discharge frequency were more strongly associated with β -sheet transformation, whereas gas flow rate showed the strongest linear association with α -helix content, and duty cycle exhibited comparatively weaker relationships. The parameter-dependent structural responses provided a basis for targeted modulation of protein secondary structures by adjusting CP processing conditions, which may help guide the development of immunoreactivity-oriented processing strategies.

3.5. Prediction of protein secondary structures using ML models

The experimental and predicted values obtained from the RF and GBDT models were presented in Fig. 7, with the corresponding model performance summarized in Table 3. Both models showed good agreement between the experimental and predicted values for the four protein secondary structure components, indicating their effectiveness in predicting changes in protein secondary structure under different CP processing conditions. Both models exhibited satisfactory predictive ability for the major secondary structural components, with relatively small differences between training and testing datasets, indicating good generalization capability and limited overfitting. For β -sheet prediction, the RF model achieved R^2 values of 0.8387 and 0.8126 for the training and testing sets, respectively, while the GBDT model showed

slightly improved performance with R^2 values of 0.8451 and 0.8195, respectively. The corresponding $RMSE$ and MAE values remained low for both models, indicating that the changes in β -sheet content induced by different CP conditions could be effectively captured. Similarly, α -helix prediction showed high accuracy, with testing R^2 values of 0.8880 for RF and 0.8902 for GBDT. These results indicated that the plasma-induced variations in ordered secondary structures exhibited relatively stable relationships with the processing parameters. Among the four secondary structural components, β -turn showed the highest prediction accuracy. The RF and GBDT models achieved testing R^2 values of 0.9075 and 0.9056, respectively, with MAE values of approximately 1.36%. The superior prediction performance of β -turn may be attributed to its strong structural coupling with β -sheet transformation, as demonstrated by the Pearson correlation analysis ($r = -0.81$). During plasma treatment, the disruption of ordered β -sheet structures was accompanied by the formation of flexible turn regions, resulting in a relatively consistent variation pattern that could be efficiently learned by ensemble algorithms. In contrast, random coil structures showed comparatively poor prediction performance, with testing R^2 values of only 0.3924 for RF and 0.3787 for GBDT. The relatively higher prediction errors ($RMSE > 1.44$) may be associated with the intrinsic structural flexibility and heterogeneity of random coil regions (Berjanskii & Wishart, 2008). Unlike α -helix and β -sheet structures, which were stabilized by relatively well-defined hydrogen-bonding patterns, random coil structures lacked a fixed spatial arrangement and may comprise multiple transient conformational states. Consequently, random coil content may show greater variability under plasma treatment, resulting in lower prediction accuracy of the ML models. Similar limitations in predicting highly flexible protein structures from spectral-derived features have also been reported in previous studies using chemometric and ML approaches (Severcan et al., 2001; Hering

et al., 2002). Comparison between RF and GBDT further revealed that both algorithms showed comparable predictive performance, although their advantages differed among secondary structures. GBDT slightly outperformed RF for β -sheet and α -helix prediction, which may be attributed to its sequential learning strategy that progressively minimized prediction errors and better captured nonlinear relationships. In contrast, RF showed marginally better performance for β -turn and random coil prediction, likely due to its strong resistance to data variability through ensemble averaging. Previous studies have shown that tree-based ensemble models can effectively capture nonlinear interactions among multiple input variables in complex biological datasets without requiring predefined mathematical relationships (Breiman, 2001; Friedman, 2002). Overall, the FTIR-ML framework quantitatively linked CP processing parameters to protein secondary-structure responses, enabling preliminary screening of candidate conditions while reducing experimental trial-and-error. Combined with subsequent SHAP analysis, the framework helped identify influential processing parameters and prioritize candidate conditions for targeted protein structural modulation. Given the potential relationship between protein conformational changes and immunoreactivity (Section 3.3), the predicted structural responses may also guide the selection of CP conditions for subsequent antibody-binding assays. Thus, this framework provides a data-driven approach for guiding CP parameter selection and targeted modulation of protein secondary structure, laying the groundwork for subsequent immunoreactivity-oriented process optimization.

3.6. SHAP interpretation of the contribution of plasma parameters to secondary structures

As shown in Fig. 8, SHAP analysis was performed for the RF and GBDT models to interpret the contributions of different plasma treatment parameters to protein

secondary structure prediction (Lundberg & Lee, 2017). The mean absolute SHAP values showed generally consistent feature rankings between the two models, indicating consistent model interpretations. Overall, discharge frequency, gas flow rate, and duty cycle showed relatively higher contributions to the predicted protein secondary structures, whereas treatment time showed relatively lower contributions. For β -sheet prediction, discharge frequency showed the highest SHAP importance in both models, followed by duty cycle and gas flow rate. The SHAP beeswarm plots (Fig. 9) further showed that high discharge frequency values (red points) were mainly distributed in the negative SHAP region, indicating that higher discharge frequency contributed negatively to the predicted β -sheet content. In comparison, the SHAP distribution of duty cycle covered a wider range, indicating that its contribution to the predicted β -sheet content varied across the input range and may reflect a nonlinear relationship. For α -helix and β -turn prediction, gas flow rate showed the highest contribution in both the RF and GBDT models. The SHAP beeswarm plots indicated that higher gas flow rates were mainly associated with positive SHAP values for α -helix prediction, suggesting a positive contribution to the predicted α -helix content. Meanwhile, gas flow rate also showed substantial contributions to β -turn prediction, with SHAP values distributed on both sides of zero, indicating that its contribution varied across the input range. This pattern may reflect the influence of gas flow rate on the transport and residence of reactive species during CP treatment rather than a simple monotonic change in protein structural alteration. For random coil prediction, duty cycle showed the greatest SHAP contribution, although its SHAP values exhibited a relatively scattered distribution. The broad distributions of SHAP values across different parameters indicated that random coil prediction was influenced by multiple processing parameters rather than a single dominant variable. This finding was

consistent with the relatively low predictive performance observed for random coil structures, which may be associated with their intrinsic structural flexibility and conformational heterogeneity. Unlike more ordered secondary structures, random coils represent dynamic ensembles of multiple transient conformations, which may increase prediction uncertainty (Craveur et al., 2015). More importantly, the SHAP analysis identified parameter-specific opportunities for targeted modulation of protein secondary structures. Discharge frequency was the most important predictor of β -sheet content, whereas gas flow rate contributed most strongly to α -helix and β -turn predictions. These findings suggested that discharge frequency could be prioritized when investigating the coordinated modulation of β -sheet and β -turn contents, while gas flow rate could be explored to regulate α -helical structure. Such structural modulation may be relevant to allergen immunoreactivity because processing-induced conformational rearrangements can alter the accessibility of antibody-binding epitopes. For example, high-intensity ultrasound and electric-field treatments modified the structure of Act d 2 and reduced its IgE-binding capacity (Wang et al., 2020; Wang et al., 2021a). Microwave processing also reduced Act d 2 content and altered protein secondary structures (Wang et al., 2023). These results indicated that changes in individual secondary structure components did not have a universally predictable relationship with immunoreactivity; rather, the effects depended on the resulting protein conformation and the accessibility of relevant epitopes. Accordingly, the SHAP-guided selection of CP conditions could facilitate targeted investigation of β -sheet/ β -turn redistribution and α -helix modulation, followed by antibody-binding assays to determine whether the predicted structural responses are accompanied by reduced immunoreactivity. The present analysis therefore provides a basis for selecting

candidate structural-modulation strategies, but direct immunological validation is required before any immunoreactivity-reduction effect can be established.

4. Conclusion

This study established an integrated FTIR and ML approach to characterize and predict CP-induced changes in protein secondary structure in kiwifruit juice. Amide I band deconvolution combined with second-derivative analysis quantified variations in α -helix, β -sheet, β -turn, and random coil structures, indicating that CP treatment was mainly associated with protein conformational rearrangement rather than substantial disruption of the peptide backbone. The structural responses of kiwifruit proteins varied with CP processing conditions. Increased treatment time and discharge frequency were associated with decreased β -sheet content and increased formation of more flexible structures, whereas gas flow rate and duty cycle produced non-monotonic changes in secondary structure. Pearson correlation and SHAP analyses further indicated that discharge frequency and gas flow rate were among the parameters most strongly associated with changes in β -sheet and α -helix contents, respectively. Furthermore, RF and GBDT models showed excellent predictive performance for protein secondary structures, particularly for β -turn and α -helix prediction. More importantly, the integrated framework can therefore support the data-driven screening of CP conditions for targeted protein structural modulation and subsequent immunoreactivity-oriented evaluation. Because changes in protein secondary structure alone cannot establish the direction or extent of immunoreactivity changes, future studies should integrate the current structural predictions with direct immunological measurements to develop and validate predictive models linking CP processing parameters, protein structural responses, and immunoreactivity. Further validation across different kiwifruit processing forms, including fresh juice, slices, and dried products, together with

595 measurements of reactive species and assessments of protein functionality, will help
596 evaluate model generalizability and identify CP processing conditions that reduce
597 immunoreactivity while preserving product quality.

598

Conflicts of interest

The authors report that there are no competing interests to declare or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by the National Key R&D Program of China (2024YFD2100602), the Research Program (QCYRCXM-2023-061) in Shaanxi Province of China, and the Research Funding (Z1013023004) from Northwest A&F University. We would like to thank the Agricultural Products Processing Laboratory at the College of Mechanical and Electronic Engineering, Northwest A&F University, for providing technical support with the use of the freeze dryer and centrifuge. We are also grateful to Dr. Yayun Hu from the College of Food Science and Engineering, Northwest A&F University, Yangling, China, for her technical assistance with FTIR analysis.

References

- An, J.-P., Li, J., Rodrigues-Stuart, K., Dewdney, M. M., Ritenour, M. A., Wang, Y., 2024. Machine learning-based metabolomics analysis reveals the early biomarkers for Diplodia stem-end rot in grapefruit caused by *Lasiodiplodia theobromae*. Postharvest Biol. Technol. 212, 112868. <https://doi.org/10.1016/j.postharvbio.2024.112868>
- Ao, J. F., Wang, S. J., Li, X. S., Law, C. L., Chen, L., 2025. Multidimensional quality assessment (nutrients and aroma) of dried kiwifruit: Unveiling the advantages of radio frequency-vacuum drying. Food Chem. 497, 146932. <https://doi.org/10.1016/j.foodchem.2025.146932>
- Ao, J. F., Zhu, H. H., Xu, T. T., Wang, S. J., Law, C. L., Chen, L., 2026. Maturity-driven changes in radio frequency-vacuum drying of kiwifruit: Mechanistic insights from dielectric properties and microstructure evolution. Food Res. Int. 241, 119685. <https://doi.org/10.1016/j.foodres.2026.119685>
- Barth, A., 2007. Infrared spectroscopy of proteins. Biochim. Biophys. Acta Bioenerg. 1767, 1073-1101. <https://doi.org/10.1016/j.bbabbio.2007.06.004>
- Bayati, M., Karalia, S., Awad, N. M. H., Goksen, G., Ktiwari, B., Mpoojary, M., 2026. Cold plasma treatment induces protein oxidation and nitration, but enhances protein digestibility. Food Chem. 514, 149081. <https://doi.org/10.1016/j.foodchem.2026.149081>
- Bergstra, J., Bengio, Y., 2012. Random search for hyper-parameter optimization. J. Mach. Learn. Res. 13, 281-305. <https://doi.org/10.1016/j.chemolab.2011.12.002>
- Berjanskii, M. V., Wishart, D. S., 2008. Application of the random coil index to studying protein flexibility. J. Biomol. NMR 40, 31-48. <https://doi.org/10.1007/s10858-007-9208-0>

638 Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5-32.
639 <https://doi.org/10.1023/A:1010933404324>

640 Byler, D. M., Susi, H., 1986. Examination of the secondary structure of proteins by
641 deconvolved FTIR spectra. *Biopolymers* 25, 469-487.
642 <https://doi.org/10.1002/bip.360250307>

643 Chen, C., Bai, T. L., Zhang, C., Wang, Z. H., Meng, X. Y., Wu, Y., Chen, H. B., Li, X.,
644 2026. Kiwifruit allergy: epidemiology, allergenic mechanisms, and processing
645 strategies for allergenicity reduction. *J. Agric. Food Chem.* 74, 19290-19306.
646 <https://doi.org/10.1021/acs.jafc.6c04110>

647 Chen, Y. Y., Wang, Z. B., 2019. End-to-end quantitative analysis modeling of near-
648 infrared spectroscopy based on convolutional neural network. *J. Chemometrics* 33,
649 e3122. <https://doi.org/10.1002/cem.3122>

650 Craveur, P., Joseph, A. P., Esque, J., Narwani, T. J., Noël, F., Shinada, N., Goguet, M.,
651 Leonard, S., Poulain, P., Bertrand, O., Faure, G., Rebehmed, J., Ghoulane, A.,
652 Swapna, L. S., Bhaskara, R. M., Barnoud, J., Téletchéa, S., Jallu, V., Cerny, J.,
653 Schneider, B., Etchebest, C., Srinivasan, N., Gelly, J. C., de Brevern, A. G., 2015.
654 Protein flexibility in the light of structural alphabets. *Front. Mol. Biosci.* 2, 20.
655 <https://doi.org/10.3389/fmolb.2015.00020>

656 Deng, M. F., Zhang, Z. Y., Ahmed, I., Yong, L., Li, Q. L., Lin, H., Li, Z. X., 2025a. A
657 review on allergens, allergenicity assessment methods, and specific allergenicity
658 assessment strategies for novel foods. *J. Agric. Food Chem.* 73, 27177-27188.
659 <https://doi.org/10.1021/acs.jafc.5c08212>

660 Deng, S., Sheng, X. W., Dong, P. F., Wang, Y. F., Wei, Q. Y., Qian, J., Wang, J., Zhang,
661 J. H., Raghavan, V., 2025b. Safety assessment of cold plasma technology in food:
662 From molecular modification to toxicological analysis. *Food Chem.* 495, 146573.

<https://doi.org/10.1016/j.foodchem.2025.146573>
 Dharini, M., Jaspi, S., Mahendran, R., 2023. Cold plasma reactive species: Generation, properties, and interaction with food biomolecules. *Food Chem.* 405, 134746. <https://doi.org/10.1016/j.foodchem.2022.134746>
 Ding, G., Jin, K., Chen, X. Y., Li, A., Guo, Z. Q., Zeng, Y. L., 2024. Non-destructive prediction of ready-to-eat kiwifruit firmness based on Fourier transform near-infrared spectroscopy. *Postharvest Biol. Technol.* 212, 112908. <https://doi.org/10.1016/j.postharvbio.2024.112908>
 Ekezie, F. G. C., Sun, D. W., Cheng, J. H., 2017. A review on recent advances in cold plasma technology for the food industry: Current applications and future trends. *Trends Food Sci. Technol.* 69, 46-58. <https://doi.org/10.1016/j.tifs.2017.08.007>
 Esteban-Sanz, C., Tobajas, A. P., Gracia, A. P., Civera, A., Sanchez, L., Arias, E., Perez, M. D., 2025. Variety, maturity stage and postharvest storage time effect on Mal d 3 allergen from apple (*Malus domestica*). *Postharvest Biol. Technol.* 225, 113512. <https://doi.org/10.1016/j.postharvbio.2025.113512>
 Friedman, J. H., 2001. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* 29, 1189-1232. <https://doi.org/10.1214/aos/1013203451>
 Friedman, J. H., 2002. Stochastic gradient boosting. *Comput. Stat. Data Anal.* 38, 367-378. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2)
 Gao, J. L., Wang, S. J., Erdogdu, F., Marra, F., Sarghini, F., Chen, L., 2025. Machine learning-driven innovations in food processing: A systematic review of applications, challenges, and future developments. *Trends Food Sci. Technol.* 164, 105204. <https://doi.org/10.1016/j.tifs.2025.105204>
 Gautam, R., Kumar, S., Upadhyayula, S., 2024. A comprehensive review on recent breakthroughs in hydrogen production from hydrogen sulfide decomposition:

688 Harnessing the power of plasma. *Renew. Sustain. Energy Rev.* 202, 114735.
689 <https://doi.org/10.1016/j.rser.2024.114735>

690 Goormaghtigh, E., Ruyschaert, J. M., Raussens, V., 2006. Evaluation of the
691 information content in infrared spectra for protein secondary structure
692 determination. *Biophys. J.* 90, 2946-2957.
693 <https://doi.org/10.1529/biophysj.105.072017>

694 Guo, X. Z., Harbers, C. J., Heuer, H. E. J. M., Min, Q. X. X., Liu, Y., Woltering, E. J.,
695 Marcelis, L. F. M., van der Sman, R. G. M., 2026. A generic scientific machine
696 learning framework for fruit and vegetable quality prediction. *Postharvest Biol.*
697 *Technol.* 234, 114103. <https://doi.org/10.1016/j.postharvbio.2025.114103>

698 Held, S., Tyl, C. E., Annor, G. A., 2019. Effect of radio frequency cold plasma treatment
699 on intermediate wheatgrass (*Thinopyrum intermedium*) flour and dough properties
700 in comparison to hard and soft wheat (*Triticum aestivum* L.). *J. Food Qual.* 2019,
701 1085172. <https://doi.org/10.1155/2019/1085172>

702 Hering, J. A., Innocent, P. R., Haris, P. I., 2002. An alternative method for rapid
703 quantification of protein secondary structure from FTIR spectra using neural
704 networks. *Spectroscopy* 16, 53-69. <https://doi.org/10.1155/2002/503989>

705 Horta, G. O., Azevedo, P. Z., de Souza, B. R., Rodrigues, S., Fernandes, F. A. N.,
706 Wischral, D., Stringheta, P. C., Martins, E., Campelo, P. H., 2026. Tailoring whey
707 protein isolate properties through controlled cold plasma processing: excitation
708 frequency, voltage and time as key variables. *J. Sci. Food Agric.* 106, 1220-1228.
709 <https://doi.org/10.1002/jsfa.70245>

710 Hsieh, K.-C., Cheng, K.-C., Lu, T.-J., Chou, Y.-J., Ting, Y. W., 2023. Effect of plasma
711 processing gas on Ara h 1 in peanut protein extract. *Food Control* 152, 109848.
712 <https://doi.org/10.1016/j.foodcont.2023.109848>

713 Hsieh, K.-C., Ting, Y. W., 2024. Atmospheric cold plasma reduces Ara h 1 antigenicity
 714 in roasted peanuts by altering the protein structure and amino acid profile. *Food*
 715 *Chem.* 441, 138115. <https://doi.org/10.1016/j.foodchem.2023.138115>

716 Kaur, S., Kumar, Y., Singh, V., Kaur, J., Panesar, P. S., 2024. Cold plasma technology:
 717 Reshaping food preservation and safety. *Food Control* 163, 110537.
 718 <https://doi.org/10.1016/j.foodcont.2024.110537>

719 Khan, M. I. H., Sablani, S. S., Nayak, R., Gu, Y. T., 2022. Machine learning-based
 720 modeling in food processing applications: State of the art. *Compr. Rev. Food Sci.*
 721 *Food Saf.* 21, 1409-1438. <https://doi.org/10.1111/1541-4337.12912>

722 Kong, J., Yu, S., 2007. Fourier transform infrared spectroscopic analysis of protein
 723 secondary structures. *Acta Biochim. Biophys. Sin.* 39, 549-559.
 724 <https://doi.org/10.1111/j.1745-7270.2007.00320.x>

725 Kumar, S., Pipliya, S., Srivastav, P. P., 2023. nth order kinetic modelling of peroxidase
 726 and polyphenol oxidase inactivation in kiwifruit juice during cold plasma and
 727 thermal treatment. *Innov. Food Sci. Emerg. Technol.* 89, 103475.
 728 <https://doi.org/10.1016/j.ifset.2023.103475>

729 Laroque, D. A., Seó, S. T., Valencia, G. A., Laurindo, J. B., Carciofi, B. A. M., 2022.
 730 Cold plasma in food processing: Design, mechanisms, and application. *J. Food*
 731 *Eng.* 312, 110748. <https://doi.org/10.1016/j.jfoodeng.2021.110748>

732 Li, B., Peng, L. X., Cao, Y. N., Liu, S. Y., Zhu, Y. C., Dou, J. G., Yang, Z., Zhou, C. G.,
 733 2024. Insights into cold plasma treatment on the cereal and legume proteins
 734 modification: principle, mechanism, and application. *Foods* 13, 1522.
 735 <https://doi.org/10.3390/foods13101522>

736 Liu, Z. W., Liu, L. J., Zhou, Y. X., Tan, Y. C., Cheng, J. H., Bekhit, A. E. D., Inam-Ur-
 737 Raheem, M., Aadil, R. M., 2021. Dielectric-barrier discharge (DBD) plasma

738 treatment reduces IgG binding capacity of β -lactoglobulin by inducing structural
 739 changes. Food Chem. 358, 129821.
 740 <https://doi.org/10.1016/j.foodchem.2021.129821>

741 Lundberg, S., Lee, S. I., 2017. *A unified approach to interpreting model predictions*.
 742 31st Conference on Neural Information Processing Systems (NIPS 2017), Long
 743 Beach, CA, USA.

744 Maddumage, R., Nieuwenhuizen, N. J., Bulley, S. M., Cooney, J. M., Green, S. A.,
 745 Atkinson, R. G., 2013. Diversity and relative levels of actinidin, kiwellin, and
 746 thaumatin-like allergens in 15 varieties of kiwifruit (Actinidia). J. Agric. Food
 747 Chem. 61, 728-739. <https://doi.org/10.1021/jf304289f>

748 Ng, S. W., Lu, P., Rulikowska, A., Boehm, D., O'Neill, G., Bourke, P., 2021. The effect
 749 of atmospheric cold plasma treatment on the antigenic properties of bovine milk
 750 casein and whey proteins. Food Chem. 342, 128283.
 751 <https://doi.org/10.1016/j.foodchem.2020.128283>

752 Olatunde, O. O., Hewage, A., Dissanayake, T., Aluko, R. E., Karaca, A. C., Shang, N.,
 753 Bandara, N., 2023. Cold atmospheric plasma-induced protein modification: Novel
 754 nonthermal processing technology to improve protein quality, functionality, and
 755 allergenicity reduction. Compr. Rev. Food Sci. Food Saf. 22, 2197-2234.
 756 <https://doi.org/10.1111/1541-4337.13144>

757 Ozturk, S., Huang, L., Hwang, C.-A., Sheen, S., 2026. Classification and detection of
 758 Salmonella, Escherichia coli O157:H7, and Listeria monocytogenes using Fourier-
 759 transform near infrared spectroscopy coupled with machine learning. Food Res.
 760 Int. 229, 118485. <https://doi.org/10.1016/j.foodres.2026.118485>

761 Panklai, T., Kumchaiseemak, N., Seelarat, W., Sangwanna, S., Chutimayanaphat, C.,
 762 Bootchanont, A., Wattanawikkam, C., Rittidach, T., Boonyawan, D., Porjai, P.,

2025. Investigating effects of air-cold plasma jet on enzymatic activity and nutritional quality attributes of Mangosteen (*Garcinia mangostana* L.) juice. *Innov. Food Sci. Emerg. Technol.* 99, 103878. <https://doi.org/10.1016/j.ifset.2024.103878>

Parakh, A., Awate, A., Barman, S. M., Kadu, R. K., Tulaskar, D. P., Kulkarni, M. B., Bhaiyya, M., 2025. Artificial intelligence and machine learning for colorimetric detections: Techniques, applications, and future prospects. *Trends Environ. Anal. Chem.* 48, e00280. <https://doi.org/10.1016/j.teac.2025.e00280>

Rahaman, T., Vasiljevic, T., Ramchandran, L., 2016. Effect of processing on conformational changes of food proteins related to allergenicity. *Trends Food Sci. Technol.* 49, 24-34. <https://doi.org/10.1016/j.tifs.2016.01.001>

Ramazzina, I., Berardinelli, A., Rizzi, F., Tappi, S., Ragni, L., Sacchetti, G., Rocculi, P., 2015. Effect of cold plasma treatment on physico-chemical parameters and antioxidant activity of minimally processed kiwifruit. *Postharvest Biol. Technol.* 107, 55-65. <https://doi.org/10.1016/j.postharvbio.2015.04.008>

Rashvand, M., Dehkharghanian, N., Nikzadfar, M., Javed, T., Luke, L. P., O'Brien, A., Shenfield, A., Genovese, F., Matera, A., Altieri, G., Di Renzo, G. C., Zhang, H., 2025. Machine learning-driven optimization for digital transformation in non-thermal food processing. *Food Bioprocess Technol.* 18, 10283-10316. <https://doi.org/10.1007/s11947-025-04078-z>

Rout, S., Panda, P. K., Dash, P., Srivastav, P. P., Hsieh, C. T., 2025. Cold plasma-induced modulation of protein and lipid macromolecules: A review. *Int. J. Mol. Sci.* 26, 1564. <https://doi.org/10.3390/ijms26041564>

Sadat, A., Joye, I. J., 2020. Peak fitting applied to Fourier transform infrared and Raman spectroscopic analysis of proteins. *Appl. Sci.* 10, 5918.

788 <https://doi.org/10.3390/app10175918>

789 Segat, A., Misra, N. N., Cullen, P. J., Innocente, N., 2015. Atmospheric pressure cold
790 plasma (ACP) treatment of whey protein isolate model solution. *Innov. Food Sci.*
791 *Emerg. Technol.* 29, 247-254. <https://doi.org/10.1016/j.ifset.2015.03.014>

792 Shah, U., Yang, M., Salvi, D., 2025. Environmental impact assessment of generating
793 cold atmospheric pressure plasma and plasma-activated water at lab scale. *Sustain.*
794 *Food Technol.* 3, 1005-1010. <https://doi.org/10.1039/d5fb00024f>

795 Severcan, M., Severcan, F., Haris, P. I., 2001. Estimation of protein secondary structure
796 from FTIR spectra using neural networks. *J. Mol. Struct.* 565, 383-387.
797 [https://doi.org/10.1016/S0022-2860\(01\)00505-1](https://doi.org/10.1016/S0022-2860(01)00505-1)

798 Tao, S. C., Wang, S. J., Bedane, T., Zhang, E. R., Chen, L. 2026. Machine learning-
799 driven parameter priority analysis for improving radio frequency heating
800 uniformity in food drying. *LWT*, 257, 119935.
801 <https://doi.org/10.1016/j.lwt.2026.119935>

802 Tolouie, H., Mohammadifar, M. A., Ghomi, H., Hashemi, M., 2018. Cold atmospheric
803 plasma manipulation of proteins in food systems. *Crit. Rev. Food Sci. Nutr.* 58,
804 2583-2597. <https://doi.org/10.1080/10408398.2017.1335689>

805 Venkataratnam, H., Cahill, O., Sarangapani, C., Cullen, P. J., Barry-Ryan, C., 2020.
806 Impact of cold plasma processing on major peanut allergens. *Sci. Rep.* 10, 17038.
807 <https://doi.org/10.1038/s41598-020-72636-w>

808 Wang, J., Vanga, S. K., McCusker, C., Raghavan, V., 2019. A comprehensive review on
809 kiwifruit allergy: Pathogenesis, diagnosis, management, and potential
810 modification of allergens through processing. *Compr. Rev. Food Sci. Food Saf.* 18,
811 500-513. <https://doi.org/10.1111/1541-4337.12426>

812 Wang, J., Vanga, S. K., Raghavan, V., 2020. Structural responses of kiwifruit allergen

813 Act d 2 to thermal and electric field stresses based on molecular dynamics
 814 simulations and experiments. Food Funct. 11, 1373-1384.
 815 <https://doi.org/10.1039/c9fo02427a>

816 Wang, J., Wang, J., Vanga, S. K., Raghavan, V., 2021a. Influence of high-intensity
 817 ultrasound on the IgE binding capacity of Act d 2 allergen, secondary structure,
 818 and In-vitro digestibility of kiwifruit proteins. Ultrason. Sonochem. 71, 105409.
 819 <https://doi.org/10.1016/j.ultsonch.2020.105409>

820 Wang, J., Zhang, L. L., Shi, J. L., Vanga, S. K., Raghavan, V., 2023. Effect of microwave
 821 processing on the nutritional properties and allergenic potential of kiwifruit. Food
 822 Chem. 401, 134189. <https://doi.org/10.1016/j.foodchem.2022.134189>

823 Wang, J., Zhou, X. Y., Li, J. Q., Pan, D. D., Du, L. H., 2024. Enhancing the
 824 functionalities of chickpea protein isolate through a combined strategy with pH-
 825 shifting and cold plasma treatment. Innov. Food Sci. Emerg. Technol. 93, 103607.
 826 <https://doi.org/10.1016/j.ifset.2024.103607>

827 Wang, S. N., Qiu, Y., Zhu, F., 2021b. Kiwifruit (*Actinidia* spp.): A review of chemical
 828 diversity and biological activities. Food Chem. 350, 128469.
 829 <https://doi.org/10.1016/j.foodchem.2020.128469>

830 Warne, G. R., Williams, P. M., Pho, H. Q., Tran, N. N., Hessel, V., Fisk, I. D., 2021.
 831 Impact of cold plasma on the biomolecules and organoleptic properties of foods:
 832 A review. J. Food Sci. 86, 3762-3777. <https://doi.org/10.1111/1750-3841.15856>

833 Yang, H. Y., Yang, S. N., Kong, J. L., Dong, A. C., Yu, S. N., 2015. Obtaining
 834 information about protein secondary structures in aqueous solution using Fourier
 835 transform IR spectroscopy. Nat. Protoc. 10, 382-396.
 836 <https://doi.org/10.1038/nprot.2015.024>

837 Yang, X., Pan, Y. Y., Zou, S., Cheng, J. H., 2025. Delivery systems of cold atmospheric

plasma in the food industry: Modes, optimization strategies, and applications.

Trends Food Sci. Technol. 156, 104831.

<https://doi.org/10.1016/j.tifs.2024.104831>

Ye, M., Zhang, Q. L., Li, H., Weng, Y. X., Wang, W. C., Qiu, X. G., 2007. Infrared spectroscopic discrimination between the loop and alpha-helices and determination of the loop diffusion kinetics by temperature-jump time-resolved infrared spectroscopy for cytochrome c. Biophys. J. 93, 2756-2766.

<https://doi.org/10.1529/biophysj.107.106799>

Yu, Y., Chai, Y. H., Yan, Y. J., Li, Z. M., Huang, Y., Chen, L., Dong, H., 2025. Near-infrared spectroscopy combined with support vector machine for the identification of Tartary buckwheat (*Fagopyrum tataricum* (L.) Gaertn) adulteration using wavelength selection algorithms. Food Chem. 463, 141548.

<https://doi.org/10.1016/j.foodchem.2024.141548>

Yue, D. H., Lin, L., Li, R. X., Zhang, Z. J., Lu, J. F., Jiang, S. T., 2025. Effect of cold plasma and ultrasonic pretreatment on drying characteristics and nutritional quality of vacuum freeze-dried kiwifruit crisps. Ultrason. Sonochem. 112, 107212.

<https://doi.org/10.1016/j.ultsonch.2024.107212>

Zhai, Y. F., Li, X. X., Sun, J. Q., Niu, L. Y., Wu, D., Xiang, Q. S., 2026. Inhibitory effects of dielectric barrier discharge plasma on growth, patulin biosynthesis and pathogenicity of *Penicillium expansum* spores as well as its application in fresh-cut kiwifruit. Postharvest Biol. Technol. 233, 114043.

<https://doi.org/10.1016/j.postharvbio.2025.114043>

Zhang, H. X., Zhao, Q. Y., Lan, T., Geng, T. H., Gao, C. X., Yuan, Q. Y., Zhang, Q. W., Xu, P. K., Sun, X. Y., Liu, X. B., Ma, T. T., 2020. Comparative analysis of physicochemical characteristics, nutritional and functional components and

863 antioxidant capacity of fifteen kiwifruit (*Actinidia*) cultivars-comparative analysis
 864 of fifteen kiwifruit (*Actinidia*) cultivars. *Foods* 9, 1267.
 865 <https://doi.org/10.3390/foods9091267>

866 Zhang, L., Boom, R. M., Ma, Y. Z., 2025. Machine learning in automated food
 867 processing: A mini review. *Annu. Rev. Food Sci. Technol.* 16, 25-37.
 868 <https://doi.org/10.1146/annurev-food-111523-122039>

869 Zhang, T. Y., Chu, J. S., Ying, Y. T., Ye, H. C., Li, M. X., Xie, Y. R., Yi, H. Z., Che, J.,
 870 Li, B. W., Zhang, B. J., Pi, X. W., 2026. Fruit allergens and their allergenicity
 871 reduction: characteristics, technologies, and prospects. *Trends Food Sci. Technol.*
 872 177, 105980. <https://doi.org/10.1016/j.tifs.2026.105980>

873 Zhou, B., Zhao, H., Yang, X., Cheng, J.-H., 2024a. Versatile dielectric barrier discharge
 874 cold plasma for safety and quality control in fruits and vegetables products:
 875 Principles, configurations and applications. *Food Res. Int.* 196, 115117.
 876 <https://doi.org/10.1016/j.foodres.2024.115117>

877 Zhou, E. N., Li, Q. Q., Zhu, D., Chen, G., Wu, L. M., 2024b. Characterization of
 878 physicochemical and immunogenic properties of allergenic proteins altered by
 879 food processing: a review. *Food Sci. Hum. Wellness* 13, 1135-1151.
 880 <https://doi.org/10.26599/FSHW.2022.9250095>

Figure caption:

Fig. 1. Cold plasma experimental setup and functions of its main components.

Fig. 2. Experimental protocol for CP processing of postharvest kiwifruit samples.

Fig. 3. FTIR spectra of postharvest kiwifruit samples treated with CP under different processing conditions. (a) treatment time, (b) discharge frequency, (c) gas flow rate, and (d) duty cycle.

Fig. 4. Second-derivative analysis of the Amide I band of postharvest kiwifruit samples under different CP processing conditions. (a) treatment time, (b) discharge frequency, (c) gas flow rate, and (d) duty cycle.

Fig. 5. Quantitative changes in the secondary structures of kiwifruit proteins under different CP processing conditions. (a) treatment time, (b) discharge frequency, (c) gas flow rate, and (d) duty cycle.

Fig. 6. Pearson correlation analysis between CP processing parameters and protein secondary structures of kiwifruit proteins.

Fig. 7. Prediction performance of ML models for kiwifruit protein secondary structures based on CP processing parameters. (a) RF model and (b) GBDT model.

Fig. 8. Feature importance of CP processing parameters based on SHAP values for predicting kiwifruit protein secondary structures. (a) RF model and (b) GBDT model.

Fig. 9. SHAP beeswarm analysis of the contributions of CP processing parameters to the prediction of kiwifruit protein secondary structures. (a) RF model and (b) GBDT model.

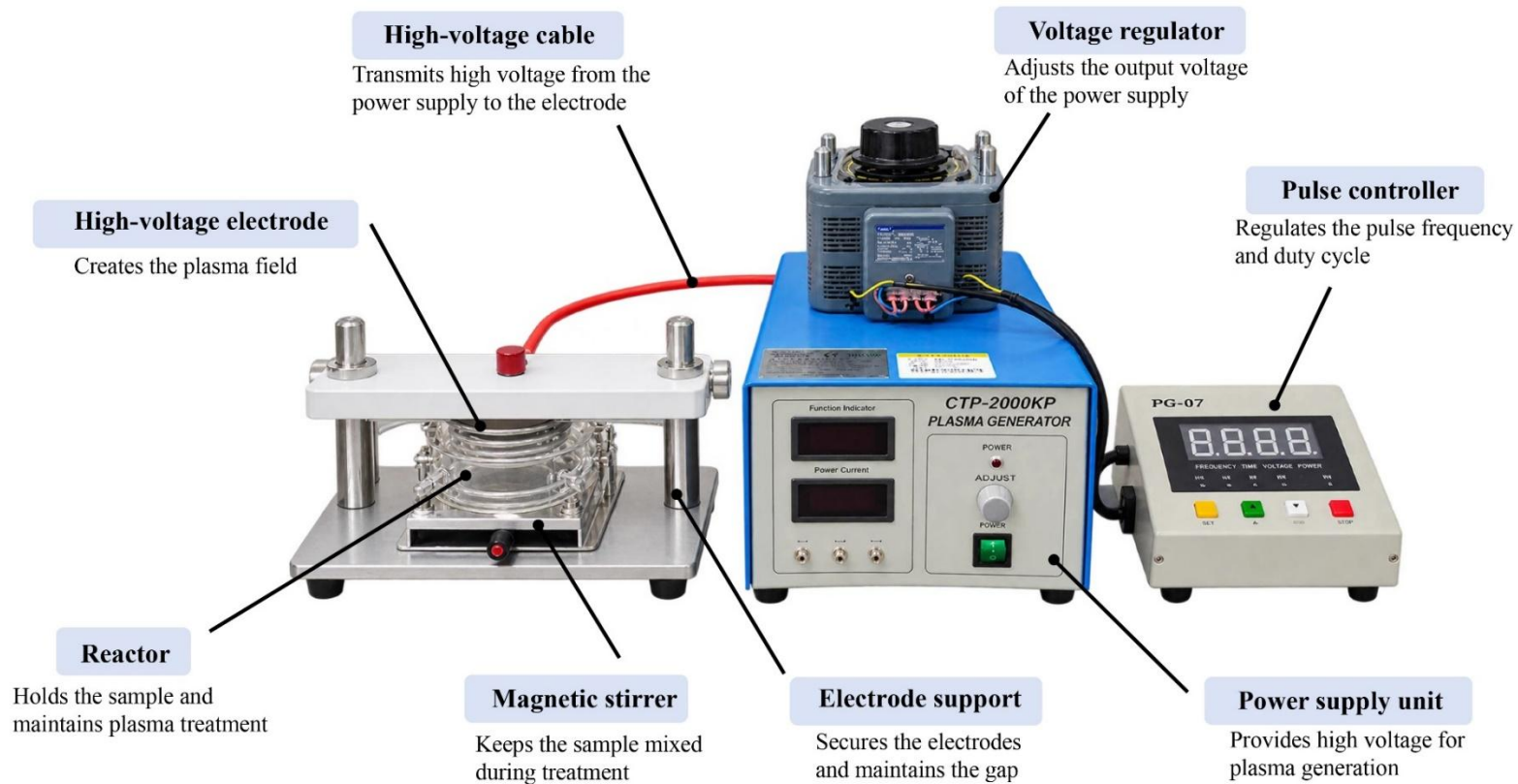


Fig. 1. Cold plasma experimental setup and functions of its main components.

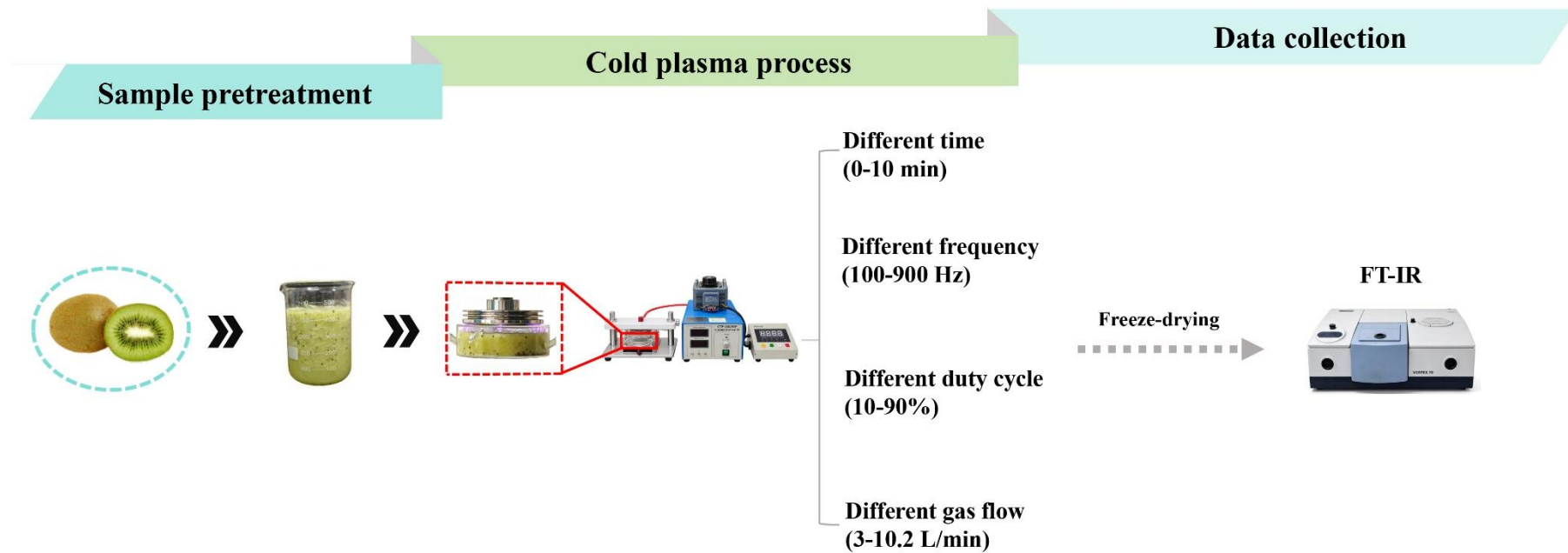
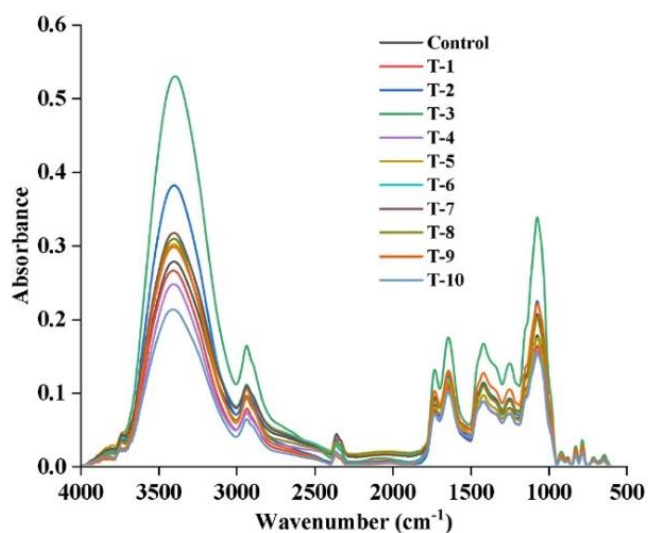
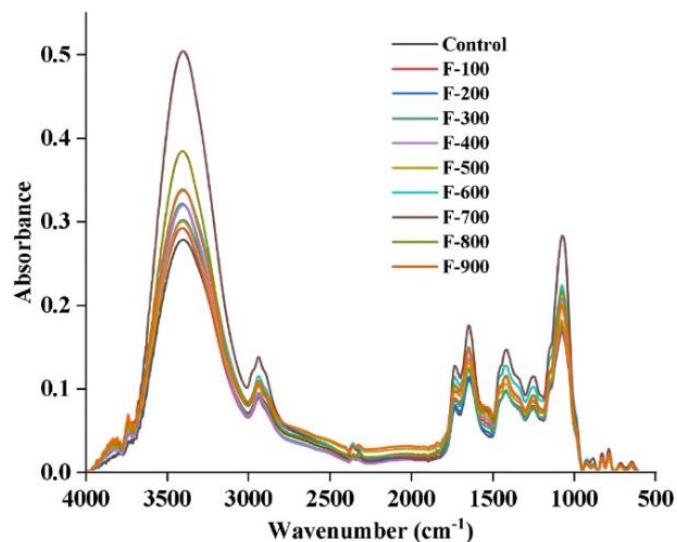


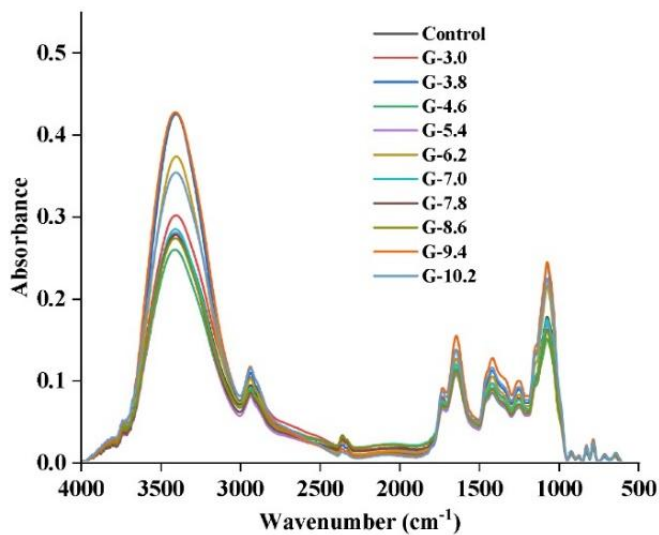
Fig. 2. Experimental protocol for CP processing of postharvest kiwifruit samples.



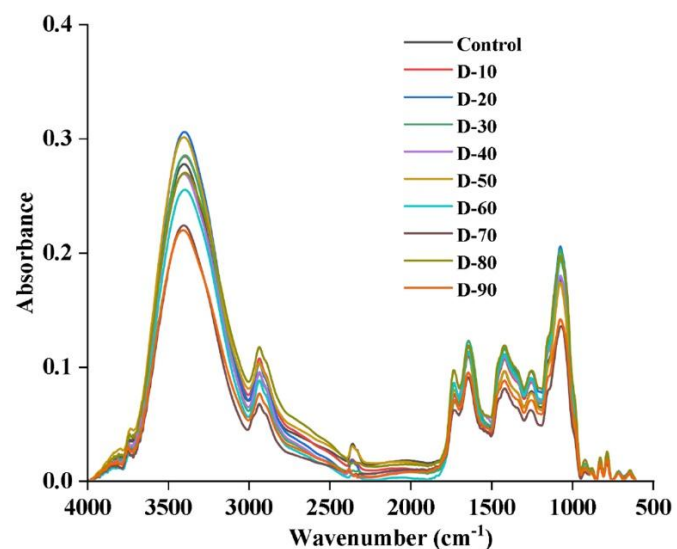
(a)



(b)

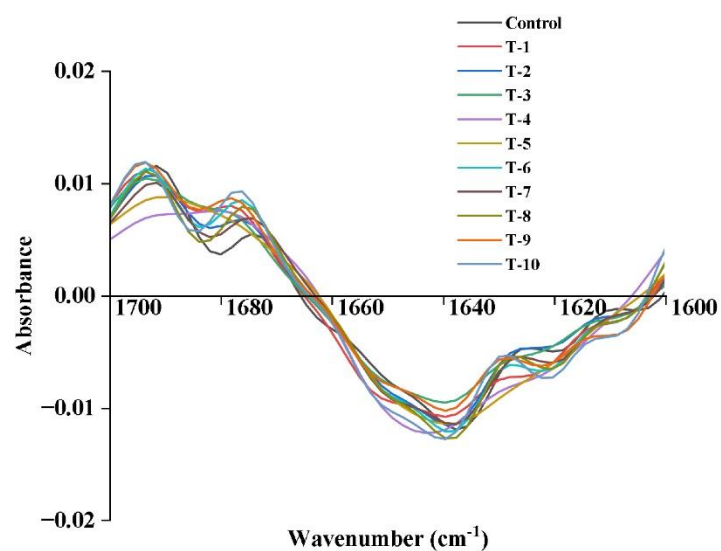


(c)

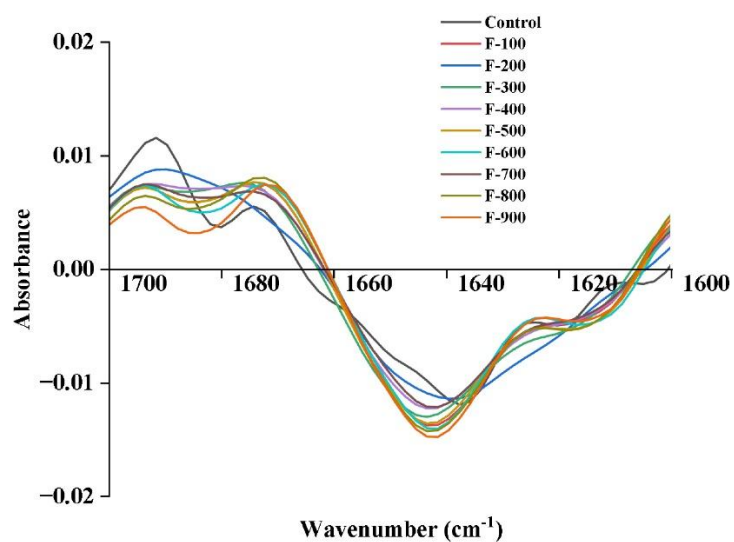


(d)

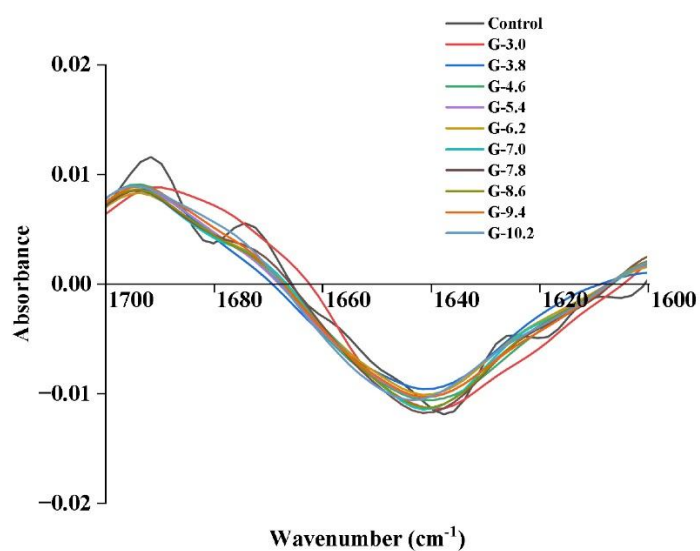
Fig. 3. FTIR spectra of postharvest kiwifruit samples treated with CP under different processing conditions. (a) treatment time, (b) discharge frequency, (c) gas flow rate, and (d) duty cycle.



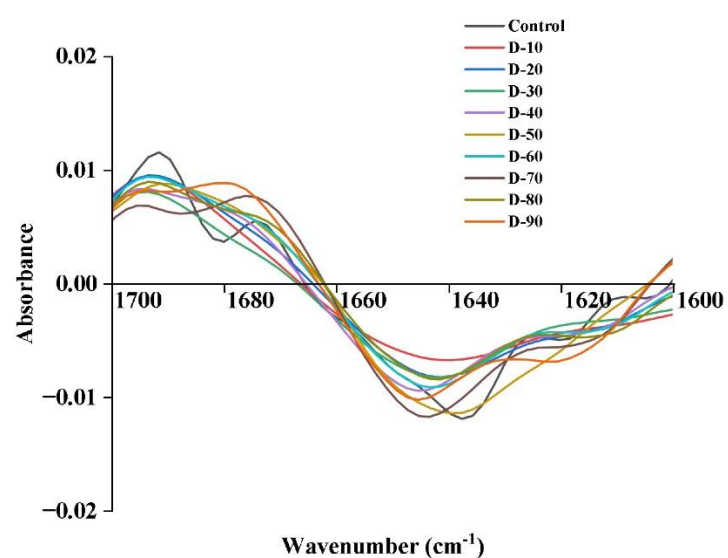
(a)



(b)

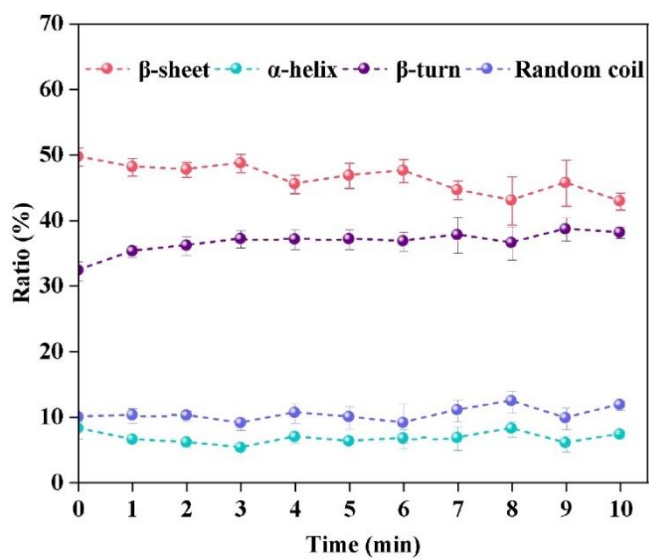


(c)

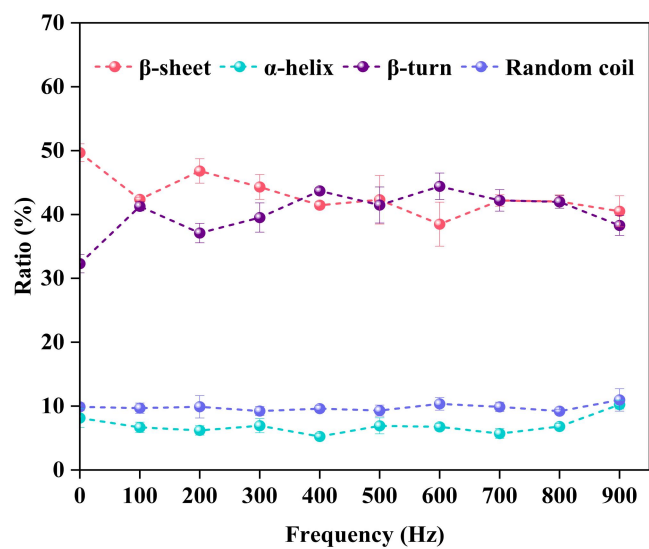


(d)

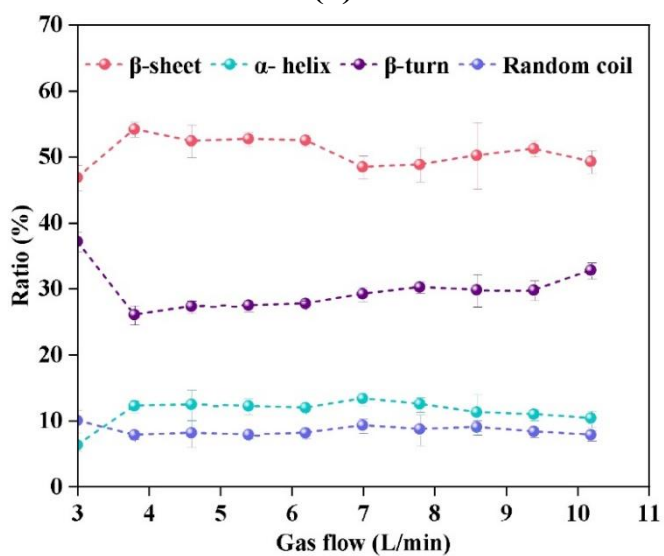
Fig. 4. Second-derivative analysis of the Amide I band of postharvest kiwifruit samples under different CP processing conditions. (a) treatment time, (b) discharge frequency, (c) gas flow rate, and (d) duty cycle.



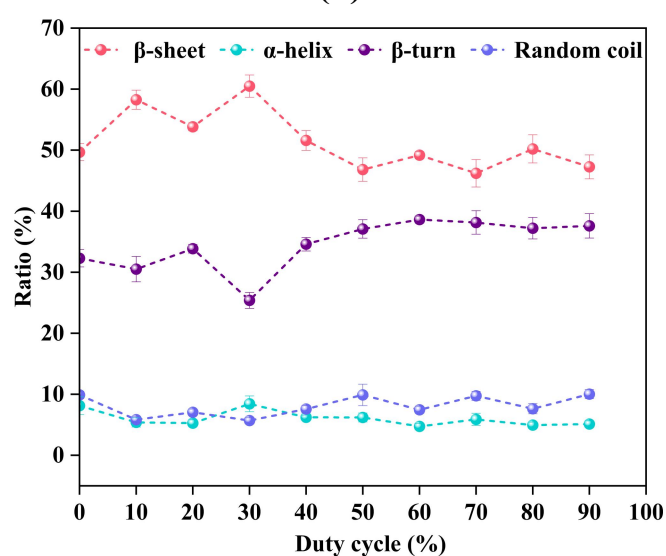
(a)



(b)



(c)



(d)

Fig. 5. Quantitative changes in the secondary structures of kiwifruit proteins under different CP processing conditions. (a) treatment time, (b) discharge frequency, (c) gas flow rate, and (d) duty cycle.

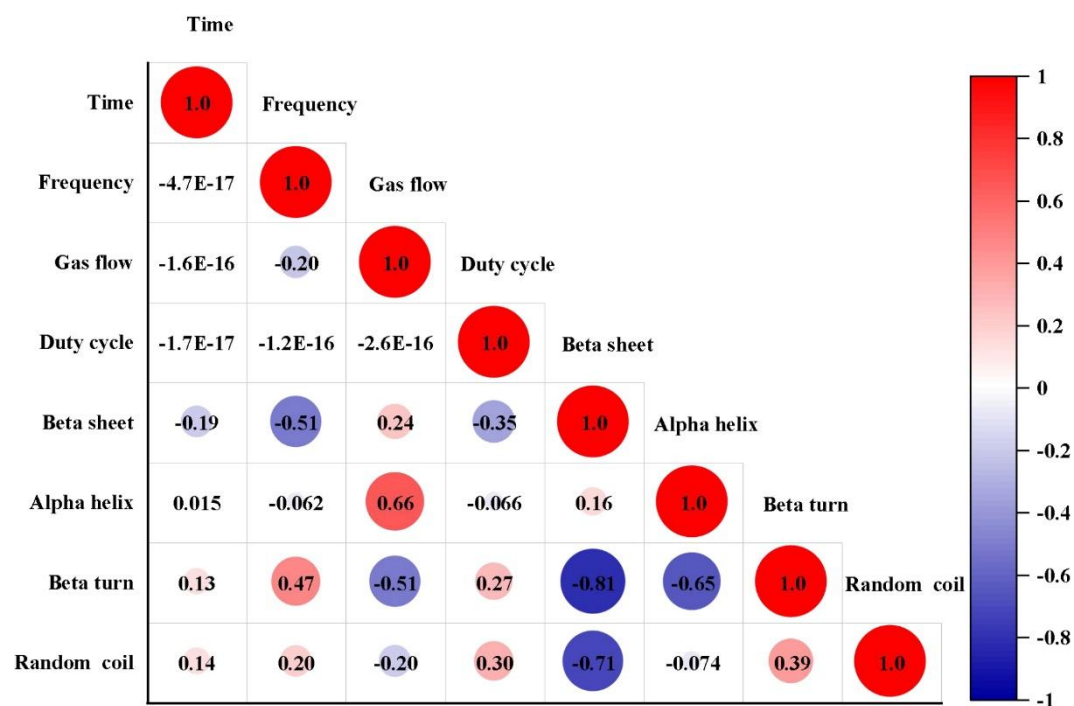
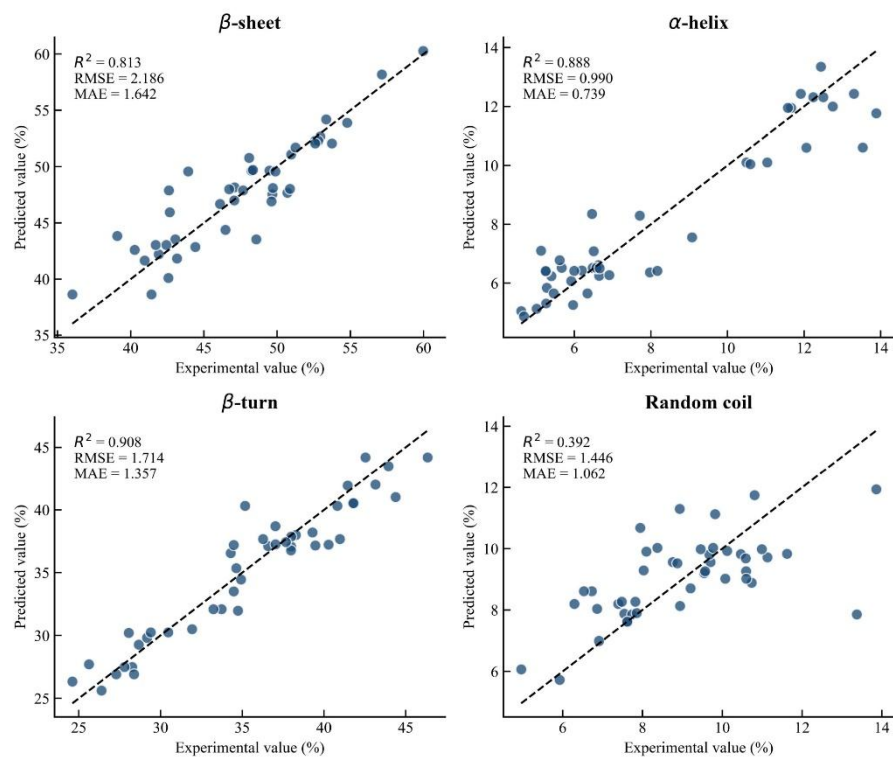
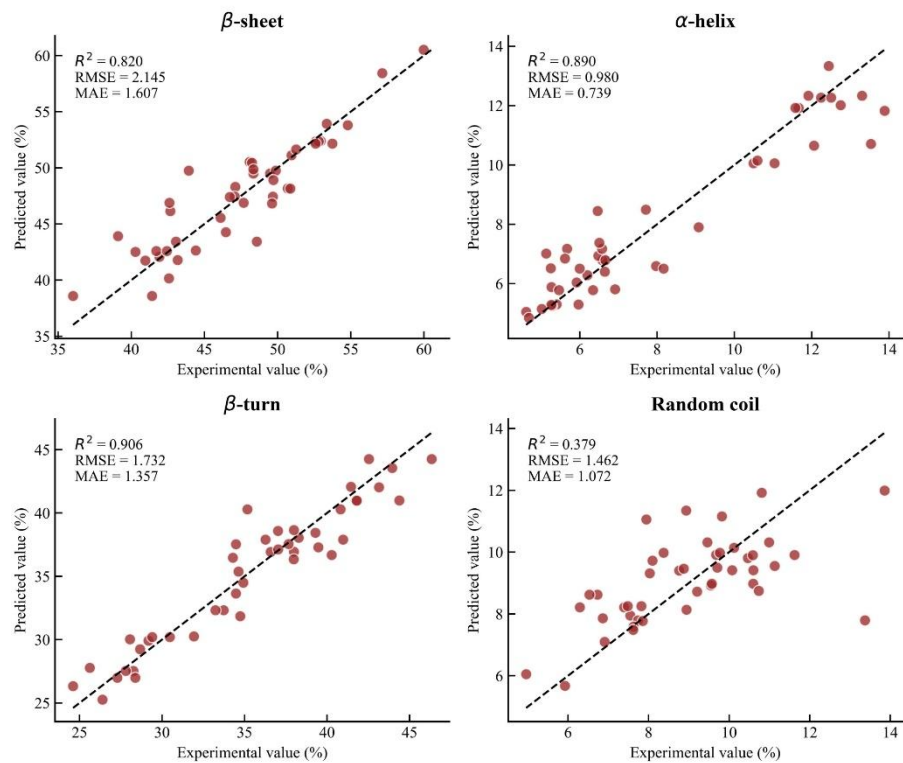


Fig. 6. Pearson correlation analysis between CP processing parameters and protein secondary structures of kiwifruit proteins.

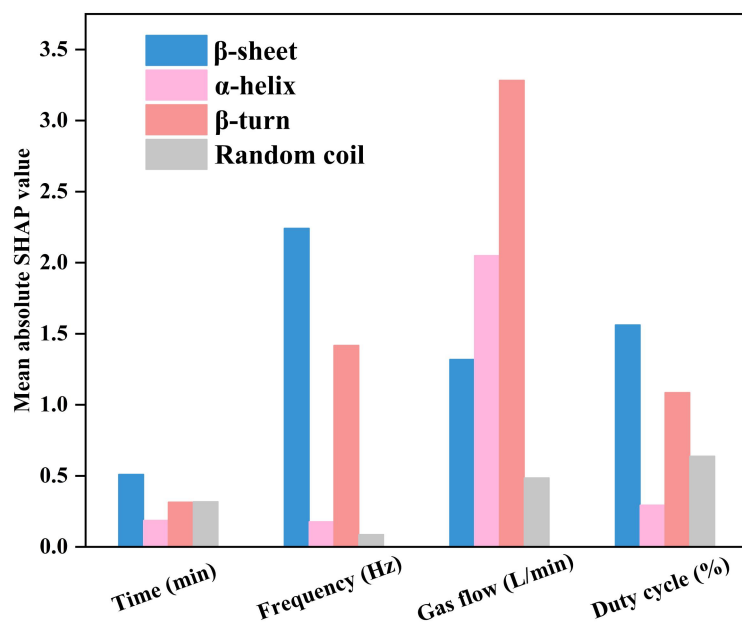


(a)

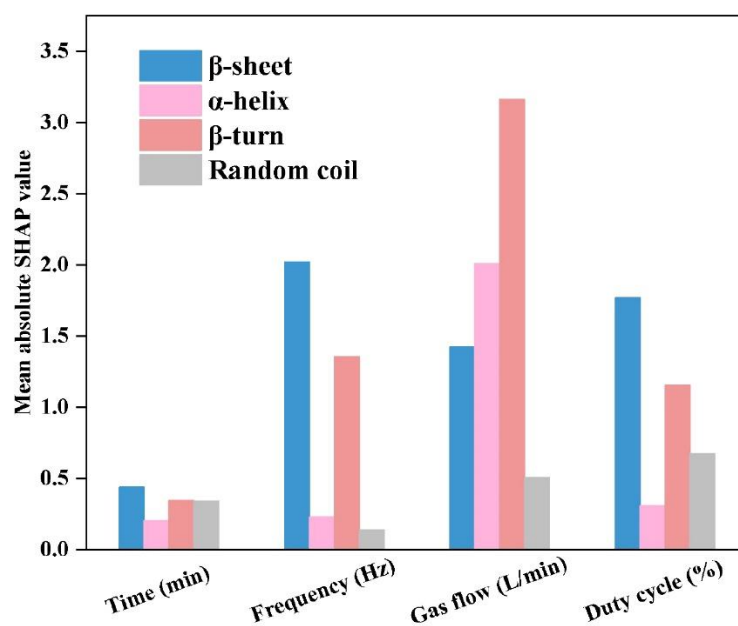


(b)

Fig. 7. Prediction performance of ML models for kiwifruit protein secondary structures based on CP processing parameters. (a) RF model and (b) GBDT model.



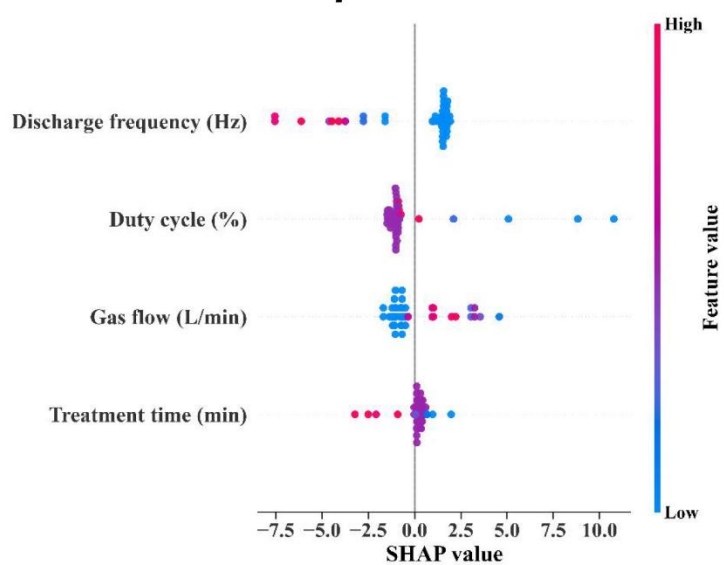
(a)



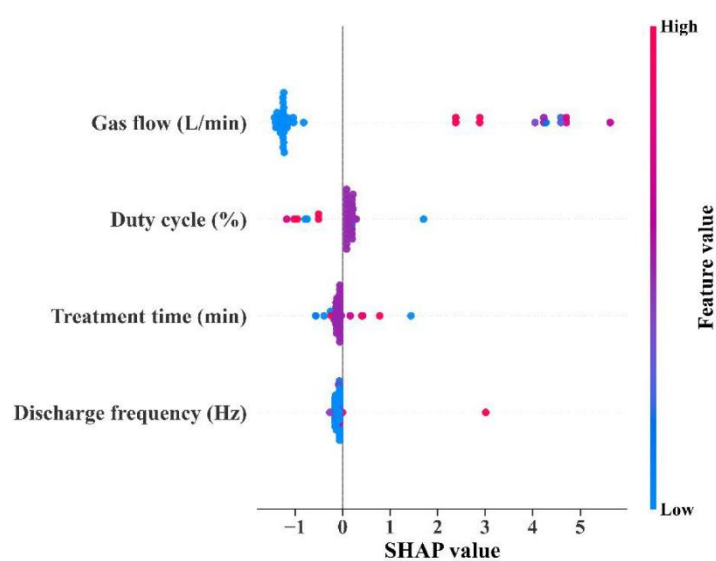
(b)

Fig. 8. Feature importance of CP processing parameters based on SHAP values for predicting kiwifruit protein secondary structures. (a) RF model and (b) GBDT model.

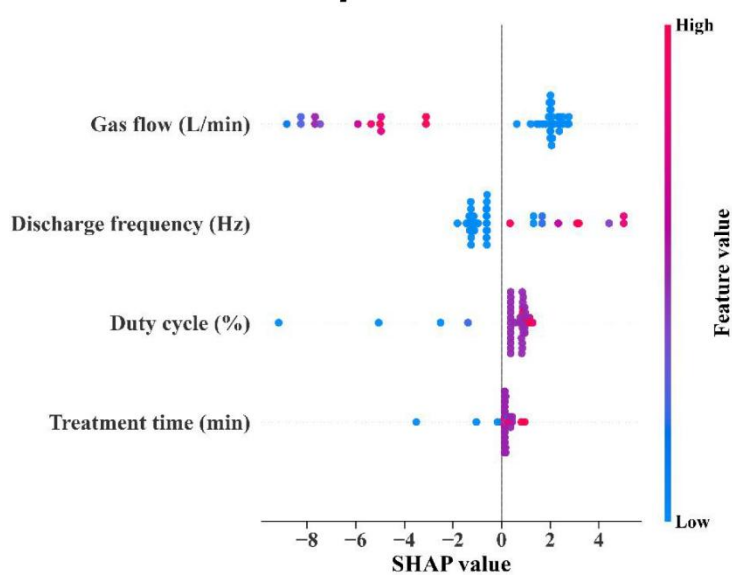
β -sheet



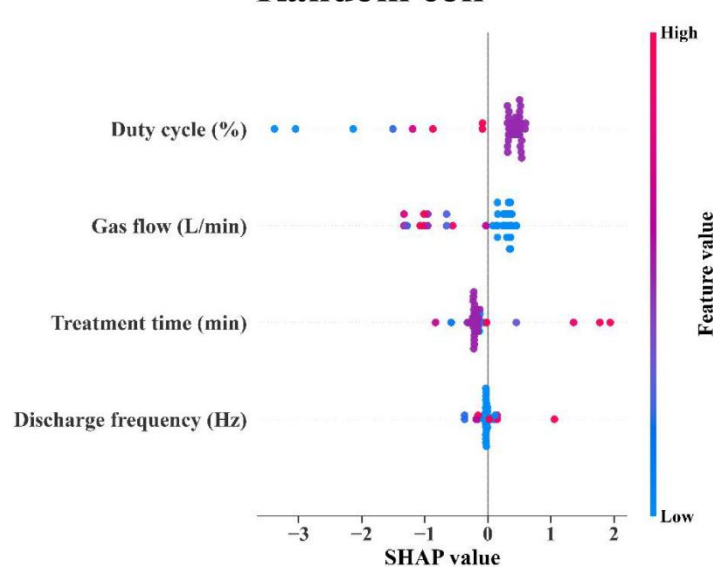
α -helix



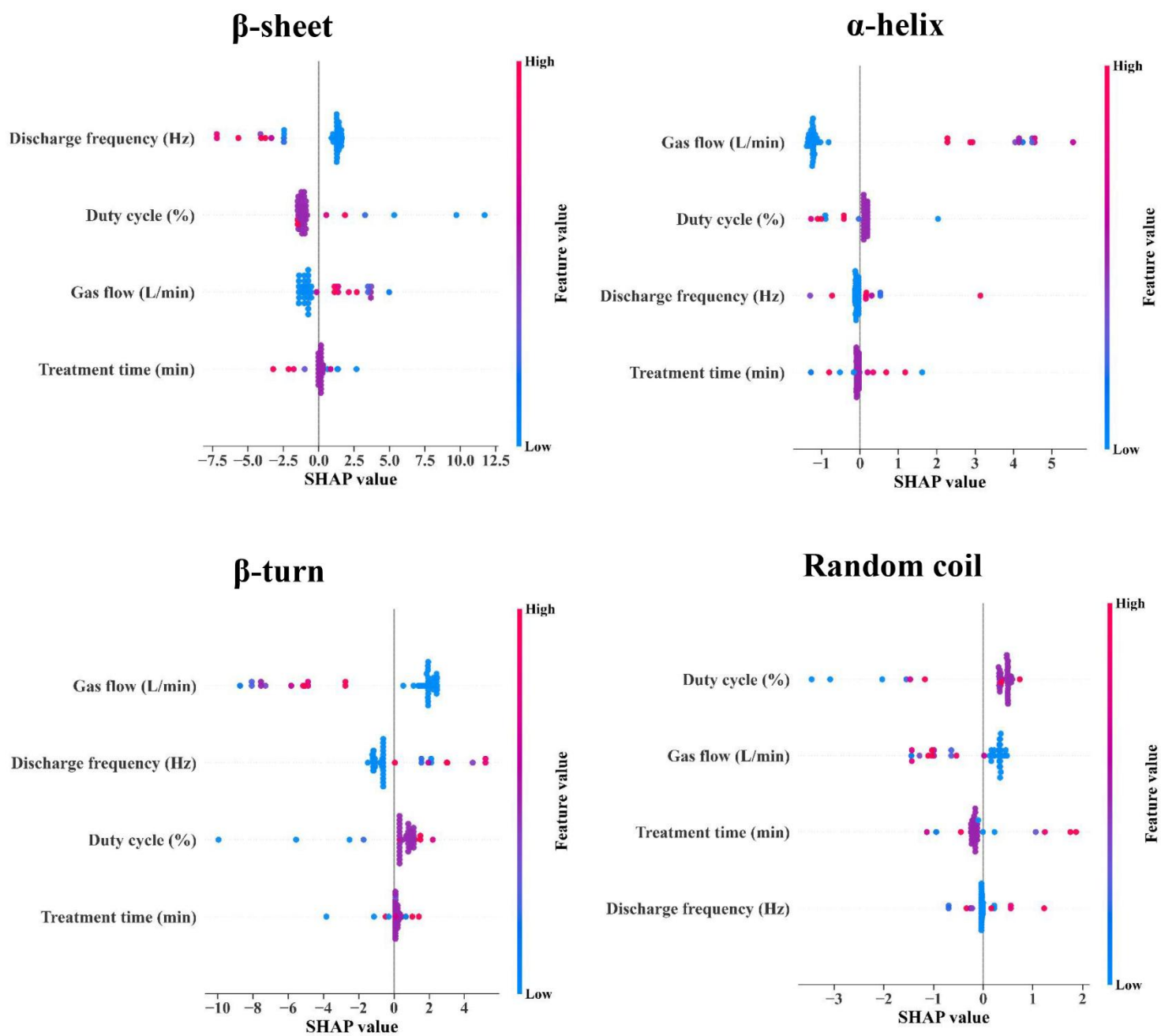
β -turn



Random coil



(a)



(b)

Fig. 9. SHAP beeswarm analysis of the contributions of CP processing parameters to the prediction of kiwifruit protein secondary structures. (a) RF model and (b) GBDT model.

Table caption:

Table 1 Curve-fitting parameters and secondary-structure assignments of the six components within the Amide I band of kiwifruit proteins.

Table 2 Relative contents of protein secondary-structure components determined by Amide I band curve-fitting analysis under different CP processing conditions.

Table 3 Performance of RF and GBDT models for predicting protein secondary-structure components based on CP processing parameters.

Table 1 Curve-fitting parameters and secondary-structure assignments of the six components within the Amide I band of kiwifruit proteins.

Peak No.	Secondary structure assignment	Wavenumber range (cm ⁻¹)	Fitted peak position (cm ⁻¹)	Peak area used for calculation	References
Peak 1	β-sheet	1615-1630	~1620	A1620	(Sadat & Joye, 2020)
Peak 2	β-sheet-related component	1630-1640	~1635	A1635	(Sadat & Joye, 2020)
Peak 3	Random coil	1640-1650	~1645	A1645	(Ye et al., 2007)
Peak 4	α-helix	1650-1660	~1655	A1655	(Goormaghtigh et al., 2006)
Peak 5	β-turn	1660-1680	~1670	A1670	(Sadat & Joye, 2020)
Peak 6	High-frequency Amide I component associated with β-turn	1680-1700	~1690	A1690	(Sadat & Joye, 2020)

Table 2 Relative contents of protein secondary-structure components determined by Amide I band curve-fitting analysis under different CP processing conditions.

Secondary structure	Included fitted components	Calculation
β -sheet	1620 and 1635 cm^{-1} components	$(A_{1620} + A_{1635})/A_{\text{total}} \times 100\%$
Random coil	1645 cm^{-1} component	$A_{1645}/A_{\text{total}} \times 100\%$
α -helix	1655 cm^{-1} component	$A_{1655}/A_{\text{total}} \times 100\%$
β -turn	1670 and 1690 cm^{-1} components	$(A_{1670} + A_{1690})/A_{\text{total}} \times 100\%$

Table 3 Performance of RF and GBDT models for predicting protein secondary-structure components based on CP processing parameters.

Model	Target structure	Dataset	R ²	RMSE	MAE
RF	β-sheet	Training	0.8387	2.1432	1.5930
		Testing	0.8126	2.1855	1.6425
	α-helix	Training	0.8201	1.1657	0.8670
		Testing	0.8880	0.9899	0.7388
	β-turn	Training	0.9096	1.5968	1.2404
		Testing	0.9075	1.7143	1.3571
	Random coil	Training	0.6257	1.1539	0.8488
		Testing	0.3924	1.4460	1.0621
	β-sheet	Training	0.8451	2.0997	1.5584
		Testing	0.8195	2.1449	1.6069
GBDT	α-helix	Training	0.8336	1.1213	0.8197
		Testing	0.8902	0.9803	0.7389
	β-turn	Training	0.9123	1.5720	1.2078
		Testing	0.9056	1.7322	1.3570
	Random coil	Training	0.6341	1.1409	0.8401
		Testing	0.3787	1.4623	1.0715